

## 基于深度学习的域适应方法综述

田青<sup>1,2</sup>, 朱雅楠<sup>1,2</sup>, 马闯<sup>1</sup>

(1. 南京信息工程大学计算机与软件学院, 南京 210044; 2. 南京信息工程大学数字取证教育部工程研究中心, 南京 210044)

**摘要:** 域适应主要应对跨不同数据分布的相似任务决策问题。作为机器学习领域的一个新兴分支, 域适应受到了众多的研究和关注。随着近年深度学习的兴起, 深度学习和域适应相结合的深度域适应研究得到了更多的关注。尽管已有各种深度域适应方法被提出, 却鲜有系统的综述工作发表。为此, 本文重点对现有的深度域适应方法进行全面回顾、分析和总结, 为相关研究人员提供借鉴和参考。本文主要贡献包括以下方面: 首先, 对域适应的背景、概念和应用领域进行概括总结。其次, 根据模型是否涉及对抗训练机制, 将现有深度域适应划分为深度对抗域适应和深度非对抗域适应两大类方法, 并逐类回顾和分析。然后, 对常用的实验基准数据集进行归类和总结。最后, 对现有深度域适应工作存在的问题和不足进行了归纳分析, 并讨论了将来的可行研究方向。

**关键词:** 域适应; 深度对抗域适应; 深度非对抗域适应; 源域; 目标域

**中图分类号:** TP391

**文献标志码:** A

## Review on Domain Adaptation Methods Based on Deep Learning

TIAN Qing<sup>1,2</sup>, ZHU Yanan<sup>1,2</sup>, MA Chuang<sup>1</sup>

(1. School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing 210044, China)

**Abstract:** Domain adaptation mainly deals with similar task decision across different data distributions. As an emerging branch of machine learning, domain adaptation has received much attention. With the rise of deep learning in recent years, the deep domain adaptation paradigm, as a combination of deep learning and traditional domain adaptation, has attracted more and more research. Although a variety of deep domain adaptation methods have been proposed, few systematic reviews have been published. To this end, this paper definitely reviews and analyzes the existing deep domain adaptation work and summarizes them to provide reference for relevant researchers. In conclusion, the main contributions of this work include the following aspects. Firstly, the background, concepts and application fields of domain adaptation are summarized. Secondly, according to whether the model training involves adversarial mechanism, we group the existing deep domain adaptation methods into two categories, such as deep adversarial domain adaptation and deep non-adversarial domain adaptation, and review and analyze them, respectively. Then,

**基金项目:** 国家自然科学基金(62176128, 61702273); 江苏省自然科学基金(BK20170956); 模式识别国家重点实验室开放课题(202000007); 模式分析与机器智能工业和信息化部重点实验室开放课题(NJ2019010); 江苏省青蓝工程项目。

**收稿日期:** 2021-06-22; **修订日期:** 2022-03-04

the benchmark datasets commonly used in the domain adaptation research are tabulated with profiles. Finally, the issues suffered in the existing deep domain adaptation work are summarized and analyzed, and future research directions are given.

**Key words:** domain adaptation; deep adversarial domain adaptation; deep non-adversarial domain adaptation; source domain; target domain

## 引言

域适应(Domain adaptation, DA),主要解决源域与目标域之间分布不同的相似任务决策问题,通过将源域知识迁移到目标域中,解决目标域任务<sup>[1-8]</sup>。常规机器学习模型通常假定训练集(源域)和测试集(目标域)来自同一分布。一旦违反这一准则,由于源域和目标域分布存在差异,此时,源域上训练的分类器应用在目标域上的效果将大打折扣。然而,在实际问题中,以上假设通常难以满足,以致预测模型在目标域上很难达到预期效果。为应对以上跨域分布差异的机器学习问题,域适应<sup>[9]</sup>的学习范式被提出。

在域适应中,通常将有监督的训练域称为源域,半监督或无监督的训练域称为目标域。根据目标域标签数据,域适应分为半监督域适应、无监督域适应和其他域适应。基于域知识迁移层面,可将域适应分为3类:基于实例加权的域适应方法<sup>[1-2]</sup>、基于特征的域适应方法<sup>[3-6]</sup>以及基于模型的域适应方法<sup>[7-9]</sup>。具体而言,基于实例加权的域适应是通过调整源域样本和类间权重以减小源域与目标域之间分布差异的;基于特征的域适应则利用新的特征表示空间的方式来拉近域之间的特征分布距离;基于模型的域适应,对源域模型知识迁移利用,以实现领域之间的对齐。以上域适应方法划分如图1所示。

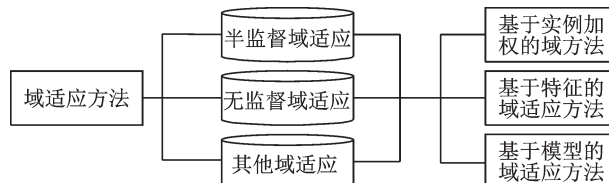


图1 域适应学习方法划分图

Fig.1 Three categories of domain adaptation methodologies

近年,以GAN<sup>[10]</sup>为代表的深度对抗学习技术开始被用于域适应学习领域。根据是否采用对抗学习的建模机制,现有的深度域适应方法可归纳为非对抗域适应和对抗域适应两类。在大多情况下,对抗域适应方法取得了总体优于非对抗域适应方法的性能。基于对抗学习机制的对抗域适应模型通常主要包括跨域表征提取模块、领域判别器等部分,通过领域判别器与跨域表征提取模块之间的对抗学习,实现领域无偏(domain-invariant)的跨域表征,最终实现领域适应。目前,对抗域适应大多采用最小-最大(min-max)的分段交替优化方式实现模型参数的求解;而部分工作通过在模型中引入梯度反转层<sup>[11]</sup>(Gradient reversal layer, GRL),进而采用梯度下降的方式实现模型求解。实际上,梯度反转层GRL作为模型训练过程中的一种优化求解方式,与分段交替优化方式在本质上相同,都是为了实现领域无偏表征的对抗生成。尽管目前已有一些域适应研究综述论文<sup>[12-16]</sup>,但鲜有系统性工作将对抗域适应方法作为单独分类进行分析总结。例如,Shao等<sup>[12]</sup>从视觉分类应用的视角回顾了域适应研究;Day等<sup>[13]</sup>对异构情景下域适应方法进行了总结。尽管以上工作<sup>[12-13]</sup>对域适应进一步发展和研究提供了借鉴意义,但并未将深度域适应方法作为回顾分析重点。除此之外,Csurka等<sup>[14]</sup>从传统建模视角罗列了一些应用较为广泛的域适应研究工作,并结合分类损失、分布差异损失和对抗机制差异对深度域适应方法进行了划分;Wang等<sup>[15]</sup>进一

步从模型应用视角出发,对文献[14]进行了扩展和补充。虽然上述工作<sup>[12-13]</sup>涵盖了部分已有对抗域适应方法,但它们侧重从其他研究视角进行分析,未能对方法的对抗建模机制进行分类和分析。鉴于对抗学习在域适应研究中的逐渐普及,本文主要基于模型结构是否涉及对抗学习机制,将现有深度域适应方法划分为深度对抗域适应和深度非对抗域适应两大类,并逐类进行回顾和分析。

如图2所示,本文着重对深度域适应方法进行回顾分析,根据建模机制将深度域适应方法分为对抗域适应和非对抗域适应两类。其中,非对抗域适应方法中包含基于差异度量的域适应、基于分布对齐的域适应和基于编码重构的域适应;对抗域适应方法包含基于标签空间的对抗域适应、基于数据增强的对抗域适应、基于距离度量的对抗域适应和基于其他机制的对抗域适应。本文主要贡献如下:首先,对域适应的背景、概念进行概括总结;其次,以是否生成对抗样本为划分标准,将深度域适应方法分为对抗域适应方法和非对抗域适应方法;然后,对域适应应用方向及常用的实验基准数据集进行归类 and 总结;最后,总结当前域适应方法研究现状,提出当前方法存在的不足以及未来可能的研究方向。

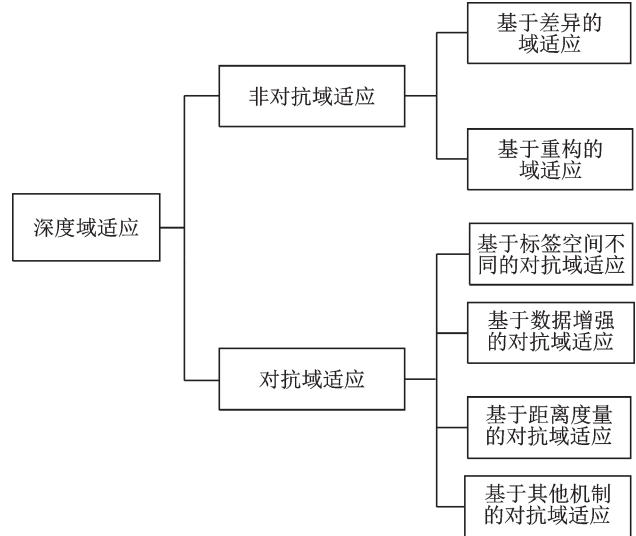


图2 深度域适应方法划分

Fig.2 Deep domain adaptation and its categories

## 1 预备知识

### 1.1 域适应概念和符号定义

域适应方法中相关概念参照文献[15]定义如下。领域(Domain)<sup>[16-17]</sup>是学习的主体,领域由数据和生成数据的概率分布这两部分组成。通常情况下, $D$ 代表Domain, $P$ 代表概率分布。在域适应中,有两类基本领域,即源域(Source domain, $D_s$ )和目标域(Target domain, $D_t$ )。源域是领域知识已知的领域;目标域是待求解的任务领域。所谓域适应<sup>[18-20]</sup>,即给定一个有标记的源域,和一个部分标记甚至无标记的目标域任务。通常而言,假定源域和目标域特征空间相同( $X_s = X_t$ ),即源域和目标域数据的维度相同( $d^s = d^t$ ),但边缘分布或条件分布不同。为后文表述方便,在此对本文所涉及的主要变量和符号进行定义,如表1所示。

### 1.2 域适应场景划分

Pan等<sup>[21]</sup>总结了不同数据集之间的差异导致领域之间的差异,即特征空间不同( $X_s = X_t$ ),或者标签空间不同( $T^s \neq T^t$ )。域适应有很多场景,不同场景对应不同的域适应方法,因此,根据是否有中间域连接,可将域适应方法分为一步域适应和多步域适应,如图3所示。根据领域差异(特征空间是否相同),可将域适应

表1 变量和符号含义

Table 1 Variables and definition of symbols

符号	含义	符号	含义
$T^{(\cdot)}$	标签空间	$x_{(\cdot)}$	样本的特征向量
$P(\cdot)$	样本分布	$y_{(\cdot)}$	样本 $x_{(\cdot)}$ 的标签值
$d^{(\cdot)}$	特征维度	$\mathcal{R}(h)$	泛化误差
$\theta^{(\cdot)}$	网络层参数	$\mu(\cdot)$	均值
$\sigma(\cdot)$	标准差	$f(\cdot)$	编码函数
$D_s$	源域样本集合	$g(\cdot)$	解码函数
$D_t$	目标域样本集合	$M(\cdot)$	关系映射

方法分为同构域适应和异构域适应。在每一类中,进一步将域适应划分为全监督域适应、半监督域适应和无监督域适应3类方法。此外,根据领域之间的标签空间大小是否相同,域适应可被划分为闭集域适应(Closed set domain adaptation)<sup>[19]</sup>、部分域适应(Partial domain adaptation)<sup>[22]</sup>、开集域适应(Open set domain adaptation)<sup>[23]</sup>和通用域适应(Universal domain adaptation)<sup>[24]</sup>。鉴于现有的绝大部分域适应研究工作研究闭集场景下的域适应问题,因此本文主要对闭集域适应<sup>[25]</sup>相关问题和工作进行总结和回顾。

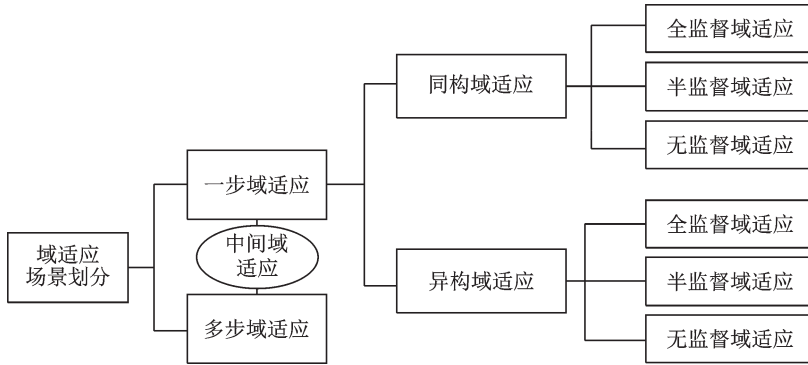


图3 域适应场景划分示意图

Fig.3 Schematic diagram of different domain adaptation scenes

### 1.2.1 同构域适应

在同构域适应中,源域和目标域有相同特征空间,即样本实例中的特征维度相同( $d^s = d^t$ ),此时,源域数据集和目标域数据集的数据分布通常不同( $P(\mathbf{X}_s) \neq P(\mathbf{X}_t)$ )。在域适应任务中,绝大多数情况下,源域样本被默认包含监督信息,即每个样本都有各自对应的标签值。极少数情况下会对源域样本是否含有数据标签进行讨论。因此本文所提出的以下分类方法中,默认所有源域样本包含标签信息,在这种情况下,本文进一步将域适应方法划分为以下3类。

#### (1) 全监督域适应

在全监督域适应中,源域提供大量有标签数据,尽管目标域的数据标签给定,但标记数量对于给定任务通常远远不够<sup>[26-27]</sup>。相较于半监督域适应和无监督域适应,全监督域适应研究工作相对较少。在全监督域适应方法中,一类典型方法是将同一类样本紧密映射到共同的潜在子空间中,同时将具有不同标签的样本分开。Saha等<sup>[26]</sup>提出的最新研究方法,采用图嵌入方式,将域标签嵌入到惩罚图结构中。

#### (2) 半监督域适应

在半监督域适应中,源域提供了大量有标签数据,目标域有充足的无标签数据和部分有标签数据<sup>[22,28-29]</sup>。具体而言,通过挖掘源域与目标域之间的相似关系,将可迁移知识从源域传递到目标域,解决目标域的分类任务。相较于无监督域适应,半监督域适应借助目标域上的少量标记信息进行域适应学习,通常能够在目标域上学得更具判别性的类别边界。因此,半监督域适应通常可获得优于无监督域适应的性能。

#### (3) 无监督域适应

在无监督域适应中,源域数据提供大量有标签数据,目标域数据的标签缺失<sup>[30-32]</sup>。因此无监督域适应没有语义信息,需要大量目标域数据才能适应源域数据分布,无监督域适应的发展最初使用MMD(Maximum mean discrepancy)方法提取域不变特征<sup>[3]</sup>。随着深度网络的兴起,无监督域适应采用神经网络与MMD结合的方式将分布对齐<sup>[30]</sup>。近几年,采用伪标签置信的方式或利用GAN网络按类别对齐的方式也成为无监督域适应研究的主流方向<sup>[32]</sup>。

### 1.2.2 异构域适应

在异构域适应中,源域和目标域的特征空间<sup>[33-36]</sup>通常不同,即样本实例中的数据维度通常不匹配( $d^s \neq d^t$ )。相较于同构域适应,尽管异构域适应的研究难度更大,但异构场景在现实生活中更加普遍。如跨语言文本分类问题<sup>[37]</sup>和跨仪器医疗影像识别问题<sup>[38]</sup>等。与同构域适应相似,异构域适应也可以根据目标域样本的标记情况,被划分为监督域适应<sup>[32-36]</sup>、半监督域适应<sup>[35]</sup>以及无监督域适应<sup>[24]</sup>3类情况。

### 1.2.3 多步域适应

在多步域适应中,知识传递通过一系列中间域实现。当源域和目标域有直接联系时,知识传递可以一步到位,通常称其为一步域适应。但实际情景中,当源域和目标域之间几乎没有重叠时,执行一步域适应方法可能导致负迁移。多步域适应大致可分为两类方法,第一类方法是中间域真实存在,如Tan等<sup>[39]</sup>提出传递迁移学习;更进一步,提出远领域迁移学习,其中源域为人脸数据集,目标域为飞机类别数据集。另一类方法是通过生成中间域作为过渡域,来解决源域和目标域之间的跨域分布差异,如Moon等<sup>[40]</sup>提出求解无监督问题的多步骤框架,并提出一种计算平均目标子空间的新方法,将计算出的变换矩阵应用于下一个目标数据作为预处理步骤,使目标数据更接近源域。Cui等<sup>[41]</sup>在生成器和判别器上均采用渐灭桥(Gradually vanishing bridge, GVB)机制,利用过渡的中间域降低整体传输难度,并减少域不变中其他域的特有特性影响。

一步域适应与多步域适应学习过程之间的差异如图4所示。具体而言,多步域适应使用一系列中间域连接两个不相关的域,通过中间域执行一步域适应,中间域能够使源域和目标域之间距离更近。

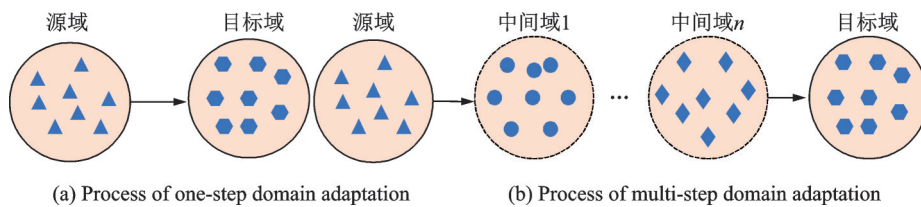


图4 一步域适应与多步域适应学习过程图

Fig.4 Learning process of diagram one-step and multi-step domain adaptations

## 2 非对抗域适应方法

深度域适应是一种利用深度网络来提高域适应性能的方法。很多深度网络模型会开发自适应层完成源域和目标域的自适应,使源域和目标域数据更加接近,从而解决目标域任务,如图5所示。深度网络中通用的网络损失定义为

$$L = l_c(x_s, y_s) + \lambda l_a(D_s, D_t) \quad (1)$$

式中: $L$ 为网络全局损失, $l_c$ 为网络的分类损失, $l_a$ 为网络自适应损失, $D_s$ 为源域数据集, $D_t$ 为目标域数据集, $x_s$ 为源域样本, $y_s$ 为源域数据集对应的标签, $\lambda$ 为权重参数。本部分将非对抗域适应方法进行划分和分析总结,主要分为基于差异度量的域适应、基于分布对齐的域适应和基于编码重构的域适应。

### 2.1 基于差异度量的域适应

在域适应中,目标域泛化误差可作为衡量目标域任务性能的标准,因此探究目标域的泛化误差是

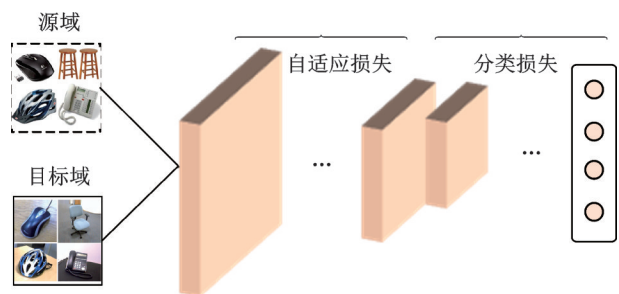


图5 深度网络中的通用损失

Fig.5 Generic losses in deep networks

极其重要的。假定  $\mathcal{H}$  为再生核希尔伯特空间。给定源域  $\mathcal{S}$  和目标域  $\mathcal{T}$ , 则有  $\forall h \in \mathcal{H}$  时

$$\mathcal{R}_{\mathcal{T}}(h) \leq \mathcal{R}_{\mathcal{S}}(h) + 1/2d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{S}, \mathcal{T}) + C \quad (2)$$

式中:  $\mathcal{R}_{\mathcal{T}}(h)$  表示目标域的泛化误差,  $\mathcal{R}_{\mathcal{S}}(h)$  表示源域的泛化误差,  $d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{S}, \mathcal{T})$  代表  $\mathcal{S}$  和  $\mathcal{T}$  两个域之间的分布差异。另外,  $C$  为最优联合泛化误差, 通常被定义为

$$C = \min_{h \in \mathcal{H}} \mathcal{R}_{\mathcal{S}}(h) + \mathcal{R}_{\mathcal{T}}(h) \quad (3)$$

由式(2)可知,  $\mathcal{R}_{\mathcal{S}}(h)$  可通过源域的监督信息最小化, 因此降低目标域泛化误差上界的关键是缩减领域之间差异  $d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{S}, \mathcal{T})$ 。根据差异度量准则的不同, 缩减领域差异的方式主要基于图准则、基于标签准则和基于架构准则等方法。

### 2.1.1 图准则

基于图准则的域适应方法, 将源域与目标域样本分别抽象为图结构。其中数据样本抽象为图结构中的顶点, 样本之间的相似度关系, 抽象为图结构中的边。因此源域与目标域数据集可被看作两个无向图。在此基础上, 如果源域数据与目标域数据所抽象的图结构相似, 则认为源域与目标域之间的差异足够小。基于以上原则, Mancini 等<sup>[42]</sup>提出了第一个基于图辅助的深度域适应框架, 并将此方法应用在域适应的连续数据场景中。该项工作通过实例化由节点所组成的图, 对源域与目标域之间的依赖关系进行建模, 本项研究为该类方法在深度域适应方向的发展奠定了良好基础。此外, Ma 等<sup>[43]</sup>从样本关系的角度出发, 采用图嵌入方法以保持不同领域样本的相似性和可辨别性。然后, 在统一的目标下, 优化矩匹配和几何对齐, 从而学习一个子空间, 以减少域间和域内的差异。另外, Morsing 等<sup>[44]</sup>也提出了双流体系结构的图嵌入域适应方法。在文中, 作者设计了两个图模型, 第一个图模型用来编码类内特征, 使同一个类内的样本足够接近。第二个图模型用来维持类间的距离, 使不同类之间的样本距离足够远。由于域适应的主要任务是找到源域和目标域的公共特征表示空间, 这一过程可被视为特征降维。而图嵌入方式是在维持图固有结构的基础上, 将图顶点映射成低维向量的过程。因此, 可用图嵌入的方式实现源域和目标域样本降维, 从而找到源域和目标域的公共特征表示空间。

上述方法, 在知识迁移和适应的过程中充分地考虑了领域之间的结构, 从而减轻了知识传递的难度。此外, 图神经网络具有良好的性能, 是一种基于深度学习的图信息处理方法, 成为了广泛应用的图分析法。迄今为止, 基于图准则的域适应方法并未和图神经网络相结合。因此在未来研究中, 可考虑图神经网络与图准则相结合的域适应方法。

### 2.1.2 标签准则

在网络分类问题中, 经常使用交叉熵损失, 交叉熵衡量的是实际输出概率与期望输出概率之间的差距, 即用来度量两个分布之间的差异。域适应中最基本的损失为分类损失, 大多数域适应方法中的分类损失采用交叉熵损失, 定义形式为

$$L = \sum_{i=1}^N \mathbf{y}^{(i)} \lg \hat{\mathbf{y}}^{(i)} + (1 - \mathbf{y}^{(i)}) \lg (1 - \hat{\mathbf{y}}^{(i)}) \quad (4)$$

式中:  $\mathbf{y}^{(i)}$  为样本真实标签,  $\hat{\mathbf{y}}^{(i)}$  为预测标签。当  $\mathbf{y}$  为 0 时, 式(4)变为  $L = -\lg(1 - \hat{\mathbf{y}})$ , 预测输出与真实输出相差越大时, 对应的惩罚项值越大, 并且成指数级增长, 这是由  $\lg$  函数本身特性所决定的。但是, 由于交叉熵损失关注于正确预测标签概率的准确性, 忽略了非正确标签的差异性, 因此利用交叉熵损失学习到的特征较为松散。Morsing 等<sup>[44]</sup>将 Softmax 输出变换为软标签输出  $\tilde{y}_i$ , 定义形式为

$$\tilde{y}_i = \frac{\exp(C_i/T)}{\sum_j \exp(C_n/T)} \quad (5)$$

式中:  $C_i, C_n$  为对应的每一类的神经元输出;  $T$  可控制标签的软化程度, 通常设为 1,  $T$  值越大意味着标签

的软化程度越高。相较于硬标签,软标签损失的不同之处在于更加注重类别之间的关系。在此启发下,Tzeng等<sup>[45]</sup>首次将软标签应用到域适应分类中,以此缩减源域与目标域之间的条件概率分布。在此之后,Timnit等<sup>[46]</sup>对软标签损失进行了改进,改进后的损失能够同时让属性级别和细粒度级别的损失达到最小。在实际问题中,目标域标签通常无法获得。因此Yan等<sup>[47]</sup>提出利用伪标签为目标域样本进行分类。一般而言,域适应中分类的准确性与置信度相关,分类的置信度越高意味着分类的结果越准确。反之,置信度越低意味着分类的结果越不可靠。而置信度由熵值衡量,其表达式为

$$w'(x) = \frac{\hat{d}' - H(\hat{y})}{\ln|C_s|} \tag{6}$$

式中: $H(\hat{y})$ 为输出的熵值; $\hat{d}'$ 为非对抗域判别器输出; $\ln|C_s|$ 为分类器的分类结果,熵值越高意味着分类的置信度越低,熵值越低意味着分类的置信度越高; $w'(x)$ 为目标域样本的可迁移性度量,是熵值和域判别输出的差与分类器分类结果的比值。虽然当前的域适应方法能够将源域丰富的标签信息转移到目标域任务中,但上述方法都基于源域标签信息不包含噪声信息这一假设。然而在现实情境下,源域的标签信息通常通过人工标注、众包收集等方式获取,不可避免地混杂错误的噪声标签。此外,目标域样本中也可能包含源域样本中不存在的无关类别标签。为应对以上问题,Yu等<sup>[48]</sup>提出了双流卷积网络框架,该网络框架结构由一个生成器和两个具有不同决策边界的分类器组成,源域的标签决策函数由双向对称的KL散度加和组成,目标域标签决策函数由KL散度和熵值加和组成。该方法能够在检测源域样本噪声标签的同时,找到目标域中的无关类别。

综上,不同的分类损失形式有各自的优缺点,分类损失作为域适应损失中较为重要的一部分,应针对实际问题和场景进行选择。

### 2.1.3 架构准则

基于架构准则的域适应方法通常分为两种情况,第一种情况通过约束网络参数,利用正则项,确保源域与目标域之间的关联性。此类方法认为,源域与目标域数据分布存在高度相似性,因此源域网络与目标域网络的参数也应该具有相关性。Rozantsev等<sup>[49]</sup>引入了双流体系结构,该结构模型中源域与目标域的对层不共享但相关,具体如图6所示。

文献[50]中假设深度学习域适应方法不应该局限于学习领域不变性特征,而应该对域迁移部分进

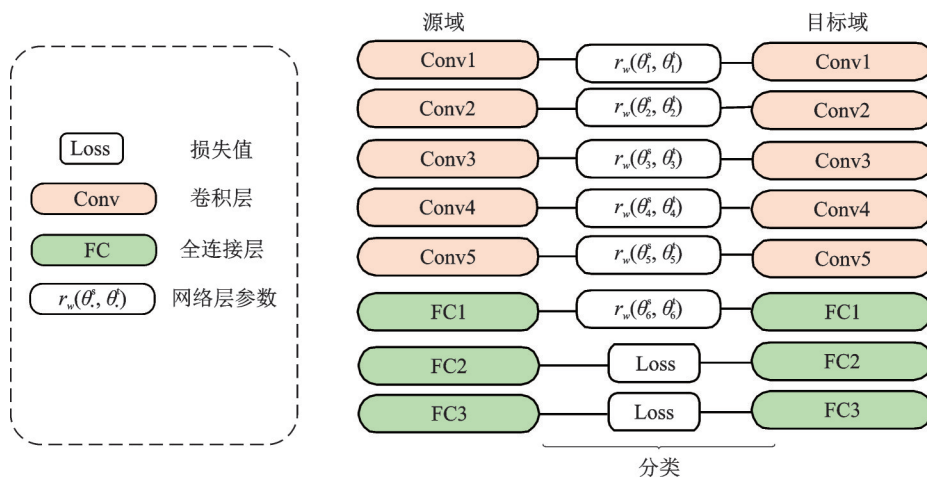


图6 基于正则项约束的双流体系网络

Fig. 6 Two-stream architecture based on regular term constraints

行显示建模。因此提出了权重正则项作为优化目标来缩减域之间差异,如式(7)所示。

$$r_w(\theta_i^s, \theta_i^t) = \exp\left(\|\theta_i^s - \theta_i^t\|^2\right) - 1 \quad (7)$$

式中: $\theta_i^s$ 为源域网络第*i*层参数, $\theta_i^t$ 为目标域网络第*i*层参数。当源域网络和目标域网络参数之差的范数 $\|\theta_i^s, \theta_i^t\|$ 为0时,指数形式 $\exp(\|\theta_i^s, \theta_i^t\|)$ 值为1,此时 $r_w(\theta_i^s, \theta_i^t)$ 达到理想情况下的最小值0,以此保证了两个网络之间的相关性。

第二种情况通过添加领域专属的批归一化层来缩减领域差异。Li等<sup>[50]</sup>假定权重矩阵中蕴含类别相关知识,批正则层蕴含域相关知识,利用BN层性质能够计算目标域均值和标准差,并将其与源域进行对齐,如式(8)所示。

$$BN(x^t) = \lambda \left( \frac{x - \mu(x^t)}{\sigma(x^t)} \right) + \beta \quad (8)$$

式中: $\lambda$ 和 $\beta$ 为目标域所学参数, $\mu(x^t)$ 为目标域网络特征通道中的均值, $\sigma(x^t)$ 为目标域网络特征通道中的标准差。式(8)中将样本值与样本均值做差后与其标准差做比值,能够保证网络层中所接收到的源域和目标域数据服从相似的分布。

总之,相较于其他方法,基于架构准则的域适应方法由最初的一个特征提取器变换为两个特征提取器,分别处理源域和目标域数据。由于源域和目标域数据分布之间存在相关性,因此两个特征提取器之间也应该存在一定的相关性。此类方法将数据之间的关系对应到模型层面,从而建立起两个领域之间的关系。

## 2.2 基于分布对齐的域适应

度量不仅是机器学习和统计学中常用的建模手段,更是域适应中常用的建模策略。度量的核心是衡量两个数据域之间的差异。具体而言,域适应中常用的分布差异度量方式包括KL(Kullback-leibler)散度、JS(Jensen-shannon)散度、沃瑟斯坦距离(Wasserstein distance)、希尔伯特施密特独立性系数、最大均值差异MMD(Maximum mean discrepancy)等。本文将分别阐述如下:

KL(Kullback-Leibler)散度<sup>[51]</sup>,又叫相对熵,本质上是两个概率分布 $P(X)$ 和 $Q(X)$ 之间的信息损失,KL散度是非对称度量,如式(9)所示。

$$D_{KL}(P||Q) = \sum_{i=1}^N P(x_i) \lg \frac{P(x_i)}{Q(x_i)} \quad (9)$$

KL散度应用在域适应问题中存在的缺点为:当两个分布无重叠时,KL散度所得到的结果可能无意义。

JS(Jensen-shannon)散度<sup>[52]</sup>则是在KL散度的基础上发展而来的,是基于KL散度的变体。解决了KL散度的非对称问题,是一种对称度量,如式(10)所示。

$$JS(P||Q) = \frac{1}{2} D_{KL}(P||E) + \frac{1}{2} D_{KL}(Q||E) \quad (10)$$

式中 $E = 1/2(P + Q)$ 。然而,JS散度具有与KL散度类似的缺点,即当分布 $P$ 和 $Q$ 距较远,完全无重叠时,JS散度值为常数,导致学习过程中梯度消失。沃瑟斯坦距离(Wasserstein distance)<sup>[53]</sup>的定义如公式(11)所示。

$$W_p(P, Q) = \left( \inf_{\mu \in W(P, Q)} \int \rho(x, y)^p d\mu(x, y) \right)^{1/p} \quad (11)$$

该距离定义在度量空间 $(E, \rho)$ 上,其中 $\rho(x, y)$ 表示集合 $E$ 中两个实例 $x$ 和 $y$ 的距离函数, $W(P, Q)$ 表示集合 $E \times E$ 内所有的以 $P$ 和 $Q$ 为边缘分布的联合分布。沃瑟斯坦距离的求解是典型的线性规划问



题,与传输矩阵中的各元素为线性关系,假设 $P(X_g)$ 和 $Q(X_f)$ 为边缘分布,则传输矩阵 $\gamma$ 为两个分布的联合分布,对应的传输矩阵示意图如图7所示。

当 $\gamma$ 边缘化 $X_g$ 和 $X_f$ 时,对应的边缘分布变为

$$P(x_g) = \int_{x_f} \gamma(x_g, x_f) dx_f \quad (12)$$

$$Q(x_f) = \int_{x_g} \gamma(x_g, x_f) dx_g \quad (13)$$

此时,利用 $\gamma$ 矩阵所得的沃瑟斯坦距离变换为

$$W(P_g, Q_f) = \inf_{\gamma \in \Pi} \iint_{x_g, x_f} \gamma(x_g, x_f) \|x_g - x_f\| dx_g dx_f \quad (14)$$

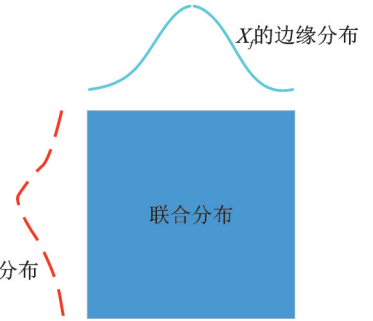


图7 传输矩阵示意图

Fig.7 Illustration of transfer matrix

沃瑟斯坦距离,能够很自然地度量离散分布和连续分布之间的距离,并且能够在保持自身分布的几何形态特征基础上,将一个分布变换为另外一个分布尽管其能够解决KL散度和JS散度的缺陷,但沃瑟斯坦距离计算量庞大。

希尔伯特-施密特独立性系数<sup>[54]</sup>本身依赖于协方差而构建,是一种基于核的度量方法。此方法在再生核希尔伯特空间中定义了互协方差算子,然后从算子中推导出度量独立性的统计量,以此来验证两组数据的独立性,如式(15)所示。

$$HSIC(X, Y) = \text{trace}(HXHY) \quad (15)$$

式中: $X, Y$ 为数据的kernel形式, $H$ 为中心矩阵。

MMD<sup>[55]</sup>度量再生核希尔伯特空间(RKHS)<sup>[56]</sup>中两个分布的距离,已被广泛应用在深度域适应中,如式(16)所示。

$$\text{MMD}^2(X, Y) = \left\| \sum_{i=1}^{n_1} \Phi(x_i) - \sum_{j=1}^{n_2} \Phi(y_j) \right\|_H^2 \quad (16)$$

式中: $\Phi(\cdot)$ 代表特征空间映射, $n_1$ 代表源域样本数量, $n_2$ 代表目标域样本数量。MMD变换过程如图8所示。

上述度量标准和相关算法总结如表2所示。

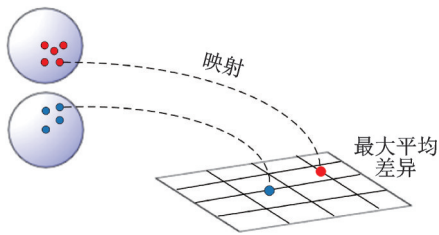


图8 MMD距离

Fig.8 MMD distance

表2 常用度量标准和相关算法

Table 2 Common metrics and related algorithms

度量准则	相关算法文献
最大均值差异(MMD)	文献[57-59]
KL(Kullback-Leibler)散度	文献[60-62]
JS(Jensen-Shannon)散度	文献[63-65]
希尔伯特-施密特独立性系数(Hilbert-Schmidt operator)	文献[66-68]
沃瑟斯坦距离(Wasserstein distance)	文献[69-71]

### 2.2.1 边缘分布对齐

按照域分布对齐方式的不同,域适应方法大致分为边缘分布域适应、条件分布域适应和联合分布域适应(图9)。边缘分布域适应利用源域和目标域边缘分布概率距离,衡量两个域之间的差异,如式(17)所示。

$$\text{Dis}_{\text{mar}}(D_s, D_t) \approx \|P(X_s) - P(X_t)\|_2^2 \quad (17)$$

式中: $\text{Dis}_{\text{mar}}(D_s, D_t)$ 代表边缘分布差异度量, $P(X_s)$ 代表源域样本分布, $P(X_t)$ 代表目标域样本分布。

Ghifary等<sup>[72]</sup>在早期提出DaNN(Domian adaptive neural networks)方法, DaNN方法仅由两层组成,

特征层和分类器层。作者衡量源域和目标域之间的损失的方法是在特征层后加入MMD适配层。但由于DaNN网络层数少,表征能力有限,因此无法有效解决域适应问题。后期很多研究工作都是在DaNN方法基础上改进。例如,加深网络层次结构,改用VGG或者AlexNet等更深的网络结构。Tzeng等<sup>[73]</sup>提出DDC(Deep domain confusion)方法训练源域和目标域的特征提取器,同时让两个网络共享参数,固定网络前7层,并在高层中加入MMD度量进行适配。

Long等<sup>[74]</sup>提出的DAN(Deep adaptation networks)方法在DDC方法上加以改进,将原来的一层适配器改为3层适配器,并且将度量改为MK-MMD(Multi Kernel-maximum mean discrepancy)。DAN方法中添加了多个核以及多个自适应层,以缩减域间分布差异。但DAN假设条件分布保持不变,这在实际应用中限制较大,且源域和目标域特征未必能通过DAN有效对齐,因此,Long等<sup>[30]</sup>提出RTN(Residual transfer networks)方法,在DAN方法基础上加以改进。DAN方法设计端到端的卷积神经网络,在原本接源域分类器的位置连接为残差网络结构,学习源域和目标域之间的差异。RTN与DAN的不同之处为,在RTN方法中需要先融合特征再对特征使用MMD惩罚。而DAN方法中直接在自适应的多个特征层中使用多个MMD惩罚。随后,Du等<sup>[75]</sup>提出了最小化梯度偏差的方法,由于目标域样本缺少标签知识,文中通过基于聚类的自监督学习方式获得目标域的可靠伪标签,进而计算目标域梯度,让源域和目标域对齐效果更佳。

该类方法主要对源域和目标域的边缘概率分布进行对齐,仅考虑了全局特征变化,忽略了类内相似性,即未能考虑源域和目标域之间的条件分布概率差异。

### 2.2.2 条件分布对齐

条件分布域适应是用源域和目标域的条件分布概率来衡量两个域之间的差异,如式(18)所示。

$$\text{Dis}_{\text{con}}(D_s, D_t) \approx \|P(Y_s|X_s) - P(Y_t|X_t)\|_2^2 \quad (18)$$

式中: $\text{Dis}_{\text{con}}(D_s, D_t)$ 代表条件分布差异度量, $P(Y_s|X_s)$ 代表了源域样本的条件分布, $P(Y_t|X_t)$ 代表了目标域样本的条件分布。

Saitod等<sup>[76]</sup>提出非对称训练方法,缩减源域与目标域之间的条件分布差异。Wang等<sup>[53]</sup>提出STL(Stratified transfer learning)方法,利用类内相似性施行类内迁移,达到更好的分类效果。Zhu等<sup>[77]</sup>提出深度子域适应网络,通过对齐特定激活层的相关子域来更新模型参数,进而使源域和目标域之间的条件分布对齐(图10)。

此类方法虽然关注了源域和目标域之间的类内相似性,却忽略了全局特征变化。目前,单纯实现条件分布对齐的域适应方法较少,其通常与边缘分布对齐联合建模使用,以达到更高的域适应分类准确性。

### 2.2.3 联合分布对齐

联合分布域适应同时考虑源域和目标域的边缘分布概率和条件分布概率,以此来衡量两个域之间的差异,如式(19)所示。

$$\text{Dis}_{\text{mar+con}}(D_s, D_t) \approx \|P(X_s) - P(X_t)\|_2^2 + \|P(Y_s|X_s) - P(Y_t|X_t)\|_2^2 \quad (19)$$

式中: $\text{Dis}_{\text{mar+con}}(D_s, D_t)$ 代表联合分布差异度量,不同于DDC、DAN、RTN等边缘分布对齐的域适应方法,以及STL等条件分布对齐的域适应方法。

Long等<sup>[78]</sup>提出的联合自适应网络JAN(Joint adaptation networks)同时从边缘分布和条件分布对齐

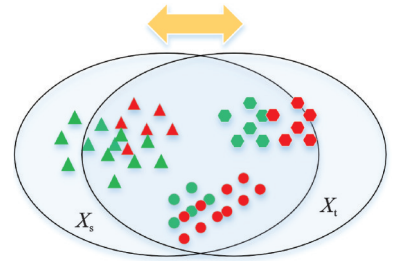


图9 边缘分布对齐图

Fig.9 Align map of edge distribution

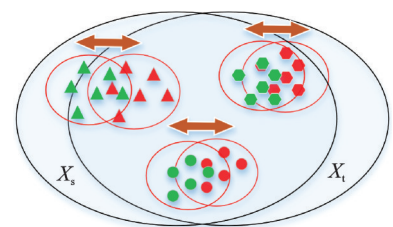


图10 条件分布对齐图

Fig.10 Align map of conditional distribution

的角度进行域适应建模。JAN 基于联合最大平均差异(JMMD)准则,在深度网络中进行联合分布自适应。此外,Zhang 等<sup>[79]</sup>提出的DTN(Deep transfer network)方法,将网络分为共享特征抽取层和判别层。样本边缘分布和条件分布基于MMD匹配,其中,共享特征提取层利用特征子空间,来对源域和目标域的数据进行边缘分布匹配,判别层利用分类器来对源域和目标域的数据进行条件分布匹配。马闯等<sup>[80]</sup>提出了一种双重加权的无监督域适应模型,首先通过目标域样本的适应性加权调整目标域样本权重,实现源域和目标域整体分布对齐,即边缘分布对齐。其次通过对源域类层面适应性加权调整源域类权重,实现源域和目标域的条件分布对齐,即条件分布对齐。

这类方法不仅关注了全局的特征变化,还利用了源域和目标域的内相似性,在对齐边缘概率分布的同时,也进行了领域之间的条件概率分布对齐(图11)。以上3种域适应方法归类总结如表3所示。

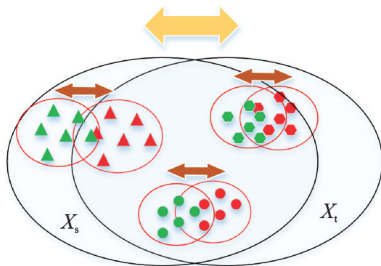


图11 联合分布对齐图

Fig.11 Align map of joint distribution

表3 3类分布对齐方式的域适应

Table 3 Domain adaptation of three types of distribution alignment

域适应方法	相关研究文献
边缘分布域适应	文献[73,81-85]
条件分布域适应	文献[69,84-85]
联合分布域适应	文献[86-89]

### 2.3 基于编码重构的域适应

基本的自动编码器框架<sup>[90]</sup>为前馈神经网络,包括编码和解码过程,通过编码函数 $f(\cdot)$ 与解码函数 $g(\cdot)$ 分别进行表示(图12)。自动编码器首先对隐藏表示的输入进行编码,然后将隐藏表示解码回重建版本,如式(20)所示。

$$z = f(x), \tilde{x} = g(z) \tag{20}$$

式中: $x$ 为输入的样本特征, $z$ 为隐含特征, $\tilde{x}$ 为重构之后的输出特征。输入的样本特征可经过编码器学习到隐含特征,学习到的隐含特征可经过解码器重构到输出特征。

Vincent 等<sup>[91]</sup>提出基于叠加去噪自动编码器(Stacked denoising autoencoders, SDA)的高级表示方法,通过同一网络,重建来自不同域的数据,高级表示可以同时表示源域和目标域数据。因此,在源域的标记数据上训练的线性分类器可以对目标域数据进行预测。尽管SDA在测试任务上取得了显著的结果,但其受计算成本高和特征维度高伸缩性差的制约。为了解决这些局限性,研究者们提出了边缘去噪自编码器(Marginalized denoising autoencoders, MSDA)<sup>[92]</sup>,MSDA用线性去噪器将噪声边缘化。用封闭形式计算参数,而且不需要随机梯度下降。MSDA 仅在20行代码中即可实现,除此之外,MSDA学到的表示与传统SDA同样有效,可在基准任务中获得与传统SDA几乎相同的精度,MSDA在域适应中的应用也取得了显著成效。

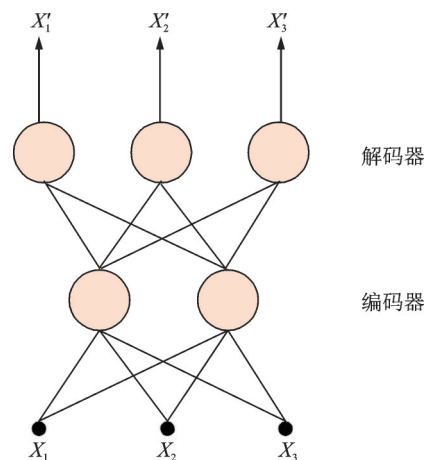


图12 自动编码器的编码与解码

Fig.12 Encoding and decoding of the autoencoder

重构域适应方法在此后得到了更深入的研究,Ghifary等<sup>[93]</sup>提出一种新的基于深度学习的无监督域适应的目标识别算法。即通过设计深度重建分类网络(Deep reconstruction classification networks, DRCN)学习共享编码表示,为跨域对象识别提供有用信息。DRCN是CNN架构,它将两条管道与共享编码器结合,如图13所示。

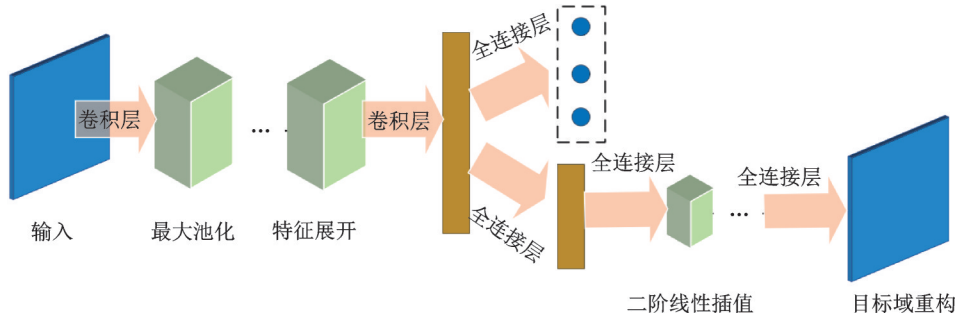


图13 深度重构分类网络框架

Fig.13 Deep reconstruction of the classification network framework

在编码器提供表示后,第1个管道(即CNN)用于具有源标签的监督分类。第2个管道(即反卷积网络)对目标域数据的无监督重建过程进行优化。DRCN为了将学习到的标签预测函数对目标域图像分类,其编码参数是共享的,而解码参数是分开的,以此将数据重建作为辅助任务以加强标签预测的准确性。

此外,域分离网络(Domain separation networks,DSNs)<sup>[94]</sup>对域表示的私有和共享组件进行建模。共享解码器通过学习私有和共享表示来重建输入样本,在共享表示上训练分类器。通过这种方式划分空间,共享表示不会受到特定于域表示的影响,可以获得更好的传输能力。模型经过训练,不仅能够源域内执行目标任务,而且可以使用分区表示从两个域中重建图像。

Tsai等<sup>[95]</sup>在分离网络和适应网络中加入混合对抗学习。文中指出,传统的域适应方法试图学习源域和目标域之间的共享表示,在没有提取源域和目标域个体信息的情况下进行分布匹配。该方法存在共享信息和个体信息混合的问题,极大限制了域适应性能。为了放松此约束,提取共享信息和个体信息至关重要。因此,该项研究通过新的域分离网络来获取共享信息和个体信息。在分离网络和自适应网络中,分别通过最小化分离损失的和最小化域偏差损失,对网络进行自适应训练。

Zhuang等<sup>[96]</sup>认为,之前的大多数工作既没有显式缩小源域与目标域之间的差异,也没有在学习表示时编码标签信息。因此笔者提出了TLDA方法(Transfer learning with deep autoencoders)。TLDA具有如下特点:(1)编码和解码权重在不同领域共享;(2)标签信息被编码。总之,基于编码重构的域适应方法是一种可抑制信息损失的无监督学习方法。该类方法通过共享编码器等方式学习跨域不变表示,并且通过源域和目标域之间的重建损失保持域私有表示。该类方法不仅可以降低信息损失,还可以将样本特征分解为域不变特征和域私有特征,从而减缓知识传递的难度。然而,该类方法通常因自编码器的表征能力有限导致其应用受限。

## 2.4 本节小结

本节对非对抗域适应方法进行了分类总结,按照实现域适应方式的不同,将其划分为基于差异度量的域适应、基于分布对齐的域适应和基于编码重构的域适应。一般而言,非对抗深度域适应方法通过深度神经网络,来减小源域与目标域的数据分布差异,进而实现域不变知识的跨域迁移和复用,并辅助目标域完成相关任务。但在实际应用中,尤其是医学图像处理上,域漂移问题仍然十分

突出,因为医学放射影像(如CT图和MRI图等)与光学仪器拍摄的自然图像的成像方式截然不同,以致这些跨模态数据之间的分布差异巨大,非对抗域适应方法在跨模态数据中的表现效果通常不佳。

### 3 对抗域适应方法

对抗域适应受GAN网络<sup>[10]</sup>的启发而近年逐渐兴起。GAN网络最初由Goodfellow等人提出,其目标是生成与训练集分布一致的数据,其中生成器负责生成样本试图混淆判别器,而鉴别器则试图区分样本的真实性,如图14所示。

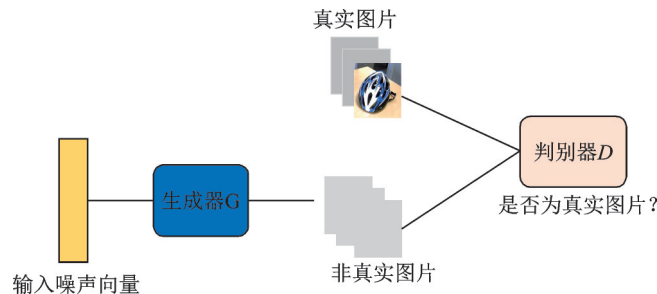


图14 GAN网络模型

Fig.14 Model of GAN

网络以对抗方式训练标签预测损失,优化 $G$ 达到最小化损失的同时训练 $D$ 最大限度地提高分配标签正确的概率,使模型最终达到纳什平衡点,GAN的总体优化函数如式(21)所示。

$$\min_G \max_D V(D, G) = E_{x \sim p(x)} [\lg D(x)] + E_{z \sim p(z)} [\lg (1 - D(G(z)))] \quad (21)$$

式中: $x \sim p(x)$ 代表 $x$ 来自真实数据分布,而 $z \sim p(z)$ 代表 $z$ 来自生成数据分布。此后,研究者们相继提出了很多GAN网络模型的变体。总体而言,分为基于结构优化的GAN网络<sup>[97-103]</sup>和基于目标函数优化的GAN网络<sup>[104-109]</sup>等,具体如图15所示。GAN网络模型的提出,使对抗学习方法成为域适应领域中常见方法之一。

基于对抗学习的域适应方法将生成对抗网络的思想引入到域适应问题中,GAN网络模型中的鉴别

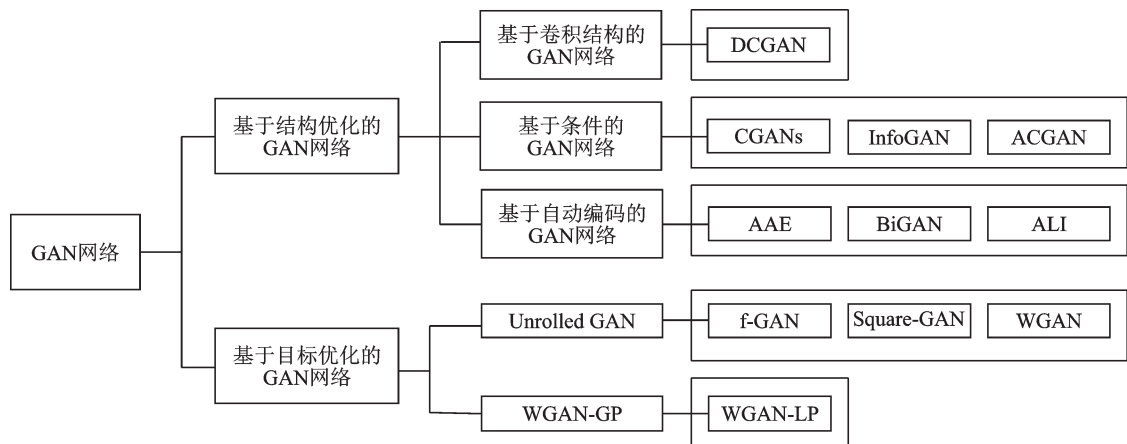


图15 GAN网络不同的衍生模型

Fig.15 Extended model from GAN

器试图准确地预测样本来自真实数据分布还是来自生成器。类似地,在域适应中可以设计鉴别器以区分源域分布和目标域分布。此时,生成器和判别器之间的博弈过程转换为特征提取与域判别之间的博弈过程。在域适应中,这一原则被用来确保网络不能区分源域和目标域样本。相较于非对抗域适应方法,对抗域适应方法能通过生成器与判别器之间的动态博弈过程更有效地提取域不变特征,并多数都取得了优于非对抗域适应方法的效果。

Tzeng等<sup>[110]</sup>提出基于对抗性方法的统一框架,并根据是否使用生成器、使用何种类别损失函数或是否跨域共享权重总结了现有的方法。Ganin等<sup>[11]</sup>首先将对抗学习应用到域适应问题当中,并且创建了DANN(Domain adversarial training of neural networks)网络,深度域适应的关键是从源域和目标域样本中学习域不变表示。从而使两个域的分布足够相似,即使在源域样本上进行训练,也可以直接用于目标域。因此,表示领域是否混淆是至关重要的转移知识,DANN所提的通用对抗域适应表达如式(22)所示。

$$E(\theta_f, \theta_y, \theta_d) = \sum_{\substack{i=1, \dots, N \\ d_i=0}} L_y^i(\theta_f, \theta_y) - \lambda \sum_{i=1, \dots, N} L_d^i(\theta_f, \theta_d) \quad (22)$$

式中: $L_y(\cdot, \cdot)$ 代表标签预测损失, $L_d(\cdot, \cdot)$ 代表域判别损失, $\theta_f, \theta_y, \theta_d$ 分别为特征映射、标签预测和域判别参数, $L_y^i(\theta_f, \theta_y) = L_y(G_y(G_f(x_i; \theta_f); \theta_y), y_i)$ 代表样本经过特征提取器之后,进入分类器进行标签预测, $L_d^i(\theta_f, \theta_d) = L_d(G_d(G_f(x_i; \theta_f); \theta_d), y_i)$ 代表样本经过特征提取器之后,进入域判别器进行域判别。在GAN网络模型的启发下,DANN引入了由鉴别器产生的域混淆损失。为了确保源域和目标域的特征分布相似,DANN将梯度反转层(GRL)加入到网络结构中。网络架构中包括特征提取器(Feature extractor)与标签预测器(Label predictor)。在反向传播过程中,DANN梯度反转层将梯度乘以固定负常数,让域分类器和特征提取器连接,实现无监督域适应过程。

Tzeng等<sup>[110]</sup>提出的对抗性判别域适应ADDA(Adversarial discriminative domain adaptation)方法通过迭代最小化3个部分的损失函数来缩小源域和目标域之间的距离,如式(23)所示。

$$\min_{M_s, C} L_{cls}(X_s, Y_s) = -E_{(x_s, y_s) \sim (X_s, Y_s)} \sum_{k=1}^K \mathbb{1}_{[k=y_s]} \lg C(M_s(X_s)) \quad (23a)$$

$$\min_D L_{advD}(X_s, X_t, M_s, M_t) = -E_{x_s \sim X_s} [\lg D(M_s(x_s))] - E_{x_t \sim X_t} [\lg(1 - D(M_t(x_t)))] \quad (23b)$$

$$\min_{M_t} L_{advM}(X_s, X_t, D) = -E_{x_t \sim X_t} [\lg D(M_t(x_t))] \quad (23c)$$

式中损失函数分为:源域经过映射后的分类损失 $L_{cls}$ ;判别器进行域判别的判别损失 $L_{advD}$ ;以及提取域不变特征损失 $L_{advM}$ 。 $M_s$ 为从源域数据中所学的映射, $M_t$ 为从目标域数据中所学的映射。相较于DANN方法,ADDA参数权重不共享,能够学习更多领域特有特征,从而更高效地缩减源域与目标域之间的差异。

DANN、ADDA等方法的提出对对抗域适应方法的发展有着重要意义,此后,很多对抗域适应方法在此基础上加以改进。本文将对抗域适应方法进行分析总结,将其划分为基于标签空间的对抗域适应、基于数据增强的对抗域适应、基于距离度量的对抗域适应和基于其他机制对抗域适应,相关总结如表4所示。

表4 四类对抗域适应建模策略

Table 4 Four kinds of adaptive modeling strategies for adversarial domains

对抗域适应方法	相关研究文献
基于标签空间的对抗域适应	文献[111-113]
基于数据增强的对抗域适应	文献[32, 114-116]
基于距离度量的对抗域适应	文献[117-118]
基于其他机制的对抗域适应	文献[119-123]

### 3.1 基于标签空间的对抗域适应

大数据时代,分类器和表示模型从现有大规模域转移到未知小规模域的情况普遍存在。因此,对抗域适应问题中存在不完全共享标签空间的情形。Cao等<sup>[111]</sup>提出SAN(Selective adversarial networks),引入标签空间不一致的对抗域适应方法,将共享标签空间这一假设放宽到:目标域标签空间是源域标签空间的一个子空间(图16)。

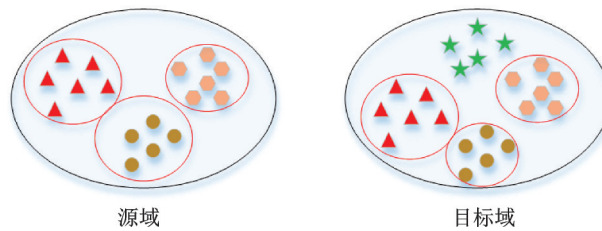


图16 标签空间不一致场景示意图

Fig.16 Illustration of inconsistent tag space

在此方法之前,通常将整个源域与目标域匹配,在标签空间不同的场景下,此类方法容易产生负迁移。因此,SAN通过筛选离群源类避免负传输,并且最大限度地匹配共享标签空间中的数据分布来促进正传输。针对标签空间的不同,Zhang等<sup>[112]</sup>也设计了一种权值机制来减轻噪声样本和离群分类的影响,并嵌入多个对抗网络来实现域之间的各类别对齐。之后,Chen等<sup>[113]</sup>提出了RTNet(Reinforced transfer network)网络结构,该项工作首次将强化学习引入到域适应问题中。文中提出,在标签空间不同的域适应问题中,大多数的方法通常为源域样本重新加权,进而选择与目标域相关的源域样本进行分布对齐。而RTNet中利用了RDS(Reinforced data selector)网络结构对离群源域样本进行过滤,其中RDS是基于Actor critic算法的强化学习网络,用来筛选相关源域样本及过滤离群源域样本。

以上方法虽然将对抗机制引入到源域与目标域分布对齐工作中,但与目标域相关的源域样本的筛选工作仍然采用其他机制进行(如权值机制和强化学习机制)。未来,能否将对抗机制引入到相关源域样本的筛选工作中是值得关注的方向。

### 3.2 基于数据增强的对抗域适应

在深度对抗域适应中,基于数据增强的方法也得到研究者的关注<sup>[114]</sup>,例如,在少数镜头学习场景中,由于缺乏数据,语义概率分布对齐和分离是困难的。因此,以CycleGAN<sup>[115]</sup>为代表的一类方法,通过最小化循环一致性损失,使得源域与目标域之间的数据能够相互转化(图17)。

在深度学习中,图像空间中的数据增强是一种成熟的技术。Chen等<sup>[116]</sup>提出像素级别的域适应方式,即利用CycleGAN进行Image-to-Image的图像增强。笔者认为,域之间进行图像转换时,预测任务应该保持一致,由此引入循环一致性损失。结果表明,相较于其他无监督域适应方法,该方法取得了更好的结果。尽管如此,但是特征空间中的数据增强未能引起同等重视。为此,Wang等<sup>[33]</sup>首次引入GAN网络模型在特征空间中进行数据扩充。他们通过GAN训练的单一特征提取器强制域不变性。进而通过定义更复杂的极大极小博弈,在特征空间中进行数据增强,论文中通过利用CGAN训练的特征生成器来实现特征空间的数据增强。其中,极大极小博弈针对特征层面进行,特征生成器允许生成符合所需类别的特征。因此CGAN生成器能够学习类在特征空间中的分布,从而生成任意数量的标记特征向量。

最近,Na等<sup>[117]</sup>通过对样本进行线性混合生成增广数据,并将增广数据作为中间域,输入到生成对抗网络中。通过在网络中进一步构建双向匹配、自惩罚和一致性正则化项约束来提升对抗建模机制的

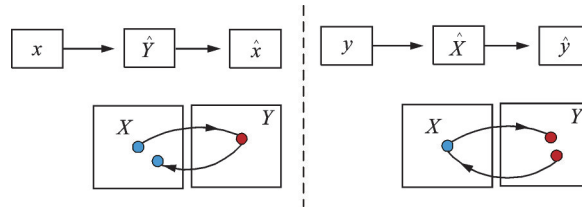


图17 循环一致性损失

Fig.17 Loss of cycle-consistency

鲁棒性。结果表明,通过特征空间数据增强,可较好地提升无监督域适应的精度。

相较于其他方法,基于数据增强的对抗域适应方法通过扩充训练样本获得了更强的模型可解释性。该类方法可对稀疏数据进行扩充,有效解决因缺少数据而导致训练模型过拟合问题。而且通过对比生成样本与真实样本差异,该类方法也可及时感知领域之间的分布差异。

### 3.3 基于距离度量的对抗域适应

部分现有研究工作没有使用鉴别器对域标签进行分类,而是在距离度量方面进行研究探索。受 Wasserstein-GAN<sup>[118]</sup> 的启发,Shen 等<sup>[119]</sup> 利用判别器估计源域样本与目标域样本之间的经验距离 (Wasserstein distance),并对特征提取网络进行了优化,以对抗的方式将距离最小化。在对抗生成网络中,当域分类器能够有效区分目标表示和源表示时,网络就会出现梯度消失的问题。文中提出更为合理的解决方案,用 Wasserstein distance 代替域差异度量,即使两个分布相距较远,也可以提供更稳定的梯度。最近,蔡德润等<sup>[120]</sup> 通过在双向长短程网络中构建注意力机制,以此作为特征提取器提取复杂的时序数据,并在域对抗方法引入加性余弦间隔损失,来解决数据整体和类别之间的分布偏移问题。

总之,度量在域适应问题中尤为重要,成为衡量领域之间差异的主要方式。然而,距离准则的优劣,通常基于模型在目标域上的性能加以判断。对于何种距离度量方式对当前域适应问题最优以及相关机理,目前仍缺少相关的理论研究。因此,如何在对抗域适应中嵌入更有效的距离度量,仍是值得关注的一个研究方向。

### 3.4 基于其他机制的对抗域适应

Rahman 等<sup>[121]</sup> 提出一种相关感知的对抗通用框架。该框架同时使用相关对齐和对抗学习来最小化源数据和目标数据特征差异。将相关对齐模块与对抗学习结合在一起,可以更有效地减少与未标记目标数据的领域差异,从而提高了模型的泛化能力。但是,丰富的领域特定特征给领域不变表示的学习带来了很大的挑战。在现有的解决方案中,领域差异被认为是直接最小化的<sup>[121]</sup>,这在实践中很难达到理想效果。部分工作通过显式地对域不变和域特定部分建模来减轻传输困难,但显式构造的不利影响在于,所构造的域不变表示中保留了域特定特征。因此,Cui 等<sup>[124]</sup> 在生成器和判别器上均采用渐灭桥 (Gradually vanishing bridge, GVB) 机制,实现了对抗域适应。GVB 不仅可以降低整体传输难度,还可以减少域不变中其他域的特有特性影响。Ahmed 等<sup>[125]</sup> 提出,由于现实中很多应用对数据安全或隐私保护的要求,领域数据有时难以获取,如医院的患者数据以及公司的客户档案等。因此,笔者提出了一种基于隐私保护的对抗域适应建模机制,通过提高伪标签置信度,增大信息熵,来联合优化源域模型对应权重,以提高目标域任务的分类精度。

当前,对抗域适应研究沿着清晰的方向在发展,基于对抗的域适应方法将对抗机制完美引入到所需解决问题的特定情景中。除上述方法外,部分方法在对抗机制中加入渐灭桥原理,通过对抗机制生



成中间域样本。此类方法在保持域不变特征的基础上模糊域私有特征,以减缓知识传递的难度。另一部分方法<sup>[126-127]</sup>采用对抗机制与注意力机制相结合的方式,选取样本图片中可迁移性强的部分进行学习。此类方法能够忽略图像中与目标任务无关的背景信息,最大程度地避免训练过程中的负迁移现象。综上所述,在未来,对抗域适应方法仍然有较大的发展空间和应用意义。

### 3.5 本节小结

本节对对抗域适应方法进行了分类总结,按照实现域适应方式的不同,将其划分为基于标签空间的对抗域适应、基于数据增强的对抗域适应、基于距离度量的对抗域适应和基于其他机制的对抗域适应。对抗域适应方法作为更具发展潜力的研究方向,拥有优于非对抗域适应方法的性能。究其原因,一方面相较于复杂的邻域间距离,对抗网络的设计相对简单,而且其提取域不变表征的能力更强。另一方面,对抗网络能够自适应实现源域和目标域之间的分布对齐,从而避免了人为的距离度量的设计。但是,受对抗网络收敛缓慢,对抗模型泛化能力不佳,容易出现模式崩塌等现象的影响,这主要是由于对抗网络模型中通常存在多个平衡态,不同平衡态之间的平衡点关联较弱所致。目前还缺少相关理论支撑推导模型的最优平衡点,这极易导致模型陷入局部最优,从而限制模型的性能。

## 4 域适应应用

域适应已广泛应用于各大领域,包括计算机视觉<sup>[122,128-139]</sup>、自然语言处理<sup>[140-146]</sup>和时间序列数据<sup>[147-152]</sup>。不同问题中利用域适应可以节省标记目标数据所需的人力时间。本文所描述的类似方法也可应用于领域泛化等相关问题。

### (1) 计算机视觉领域

现有的域适应方法大多面向计算机视觉任务,例如,将合成图像训练的模型适应于真实照片。在游戏行业中,游戏图像与真实场景之间存在领域差异,域适应成为解决该问题的主要手段。医学图像领域中,域适应也能将心脏核磁共振图转化为CT图<sup>[153-154]</sup>。此外,域适应方法也用于自主导航<sup>[122]</sup>、胸部X射线分割<sup>[128]</sup>(图18)、3D CT扫描到X射线分割<sup>[129]</sup>、目标检测<sup>[130-132]</sup>、行人重识别<sup>[133-139]</sup>等。

### (2) 自然语言处理领域

域适应还可用于自然语言处理问题,如情绪分析<sup>[140]</sup>、文本分类<sup>[140,142]</sup>、问题回答<sup>[143]</sup>、机器翻译<sup>[144-146]</sup>、词表征学习<sup>[155]</sup>以及关系提取<sup>[156]</sup>等。基于深度学习的域适应方法使得自然语言处理众多任务取得了重大进展。一方面,可以成功运用于分词、词性标注、成分分析、语义表示等任务;另一方面,可以提升问答系统、对话系统、阅读理解等应用系统的性能。

### (3) 时间序列数据领域

域适应在时间序列数据中应用广泛,既可用于学习提取不同年龄人口健康数据中的时间潜在关系<sup>[147]</sup>,也可用于语音识别<sup>[148]</sup>、预测驾驶动作<sup>[149]</sup>、异常检测<sup>[150]</sup>等,其中,域适应在异常检测中的应用如图19所示。

在领域泛化相关问题方法中,时间序列无线电数据被用于睡眠阶段分类<sup>[151]</sup>,时间序列数据也可以结合预训练和微调等方式,解决源数据集与目标数据集标签空间不同的问题<sup>[152]</sup>。

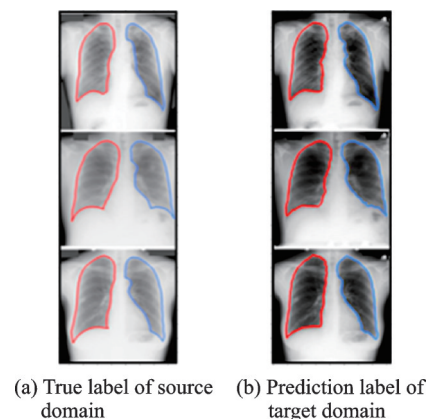


图18 胸部X射线分割图

Fig.18 X-ray segmentation of the chest

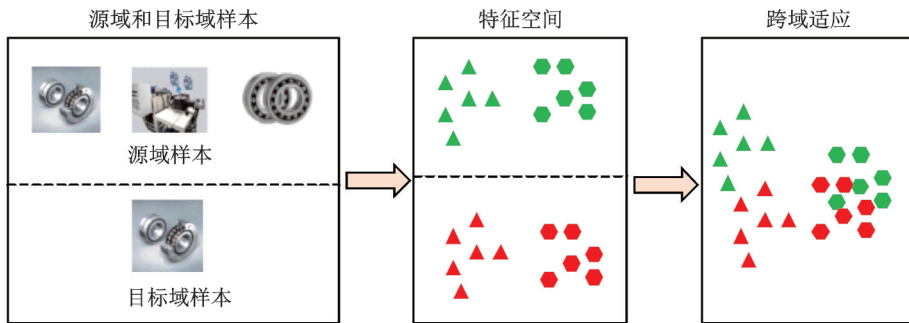


图19 域适应在异常检测中的应用

Fig.19 Application of domain adaptation in anomaly detection

(4) 域泛化领域

域不变特征学习方法已经用于域泛化相关问题,即存在多个源域和一个未知目标域<sup>[157-158]</sup>。Ghifary等<sup>[159]</sup>使用去噪自动编码器,利用重建方法提高物体识别的准确性。其中“噪声”是数据的不同视图(域)(例如,旋转、大小变化,或照明变化)和自动编码器试图重建在其他领域对象的相应视图。Carlucci等<sup>[160]</sup>在进行域映射的同时,提出了域适应和域泛化相结合的对抗性方法。由于域不变表示会与区分表示相竞争,Akuzawa等<sup>[161]</sup>提出对抗特征学习与准确性约束(AFLAC)相结合的方法来找到不影响分类性能的域不变表示。

5 相关数据集与实验

相关数据集与实验结果分别如表5~8所示。

表5 实验所需数据集

Table 5 Datasets required for experiments

数据集	介绍
MNIST <sup>[162]</sup>	MNIST数据集是一个手写数字数据集(0~9),包含60 000个示例的训练集和10 000个示例测试集。是NIST集合的子集,该数据集中的图像已经被固定化大小并且进行了归一化居中处理。
MNIST-M <sup>[135]</sup>	MNIST-M数据集,由MNIST数字与BSDS500数据集中的随机色块混合而成。
USPS <sup>[163]</sup>	USPS同样为手写数字数据集(0~9),包含了7 291训练集和2 007测试集,数据集被大小归一化为16像素×16像素的灰度图。
SVHN <sup>[164]</sup>	SVHN是一个真实世界的图像数据集,用于开发机器学习和目标识别算法,对数据预处理和格式要求较低。在风格上与MNIST相似,但SVHN数据集包含了一个数量级的标记数据(超过60万数字图像),并且SVHN数据集是来自现实世界的数据集(在自然场景图像中识别数字和数字)。
SYN <sub>N</sub> <sup>[135]</sup>	SYN <sub>N</sub> 数据集使用Microsoft Windows字体创建一个合成数字数据集,该数据集由1~3位数字组成,具有不同的位置、方向、背景颜色和笔画颜色。
SYN <sub>S</sub> <sup>[163]</sup>	SYN <sub>S</sub> 数据集是一个合成的标志数据集,由维基百科上的交通标志象形图修改而成。它由10万幅图像和43类标志组成。
GTSRB <sup>[164]</sup>	GTSRB数据集是一个德国交通标志检测数据集,包含约5万幅图像和43种标志。
Office <sup>[165]</sup>	Office数据集包含Amazon、Webcam和Dslr数据集,在每个场景包含31个类别。
ImageCLEF-DA <sup>[166]</sup>	ImageCLEF-DA数据集是域适应任务中的基准数据集之一,共包含3大数据集。其中每个数据集中包含12个类别,每个类别中包含50张图像。
VisDA-2017 <sup>[166]</sup>	VisDA-2017数据集是无监督域适应任务中的大型基准数据集之一,包含图像分类和图像分割这两大任务。共含有12个类别和28万张图像。
DomainNet <sup>[167]</sup>	DomainNet数据集是多源域适应任务的大型基准数据集之一,共包含六大数据集。其中每个数据集中包含345个类。

表6 Office31数据集上的评价结果

Table 6 Evaluation results on the Office31 dataset

评价算法	A-W	D-W	W-D	A-D	D-A	W-A	平均
ResNet-50 <sup>[168]</sup>	68.4	96.7	99.3	68.9	62.5	60.7	76.1
DAN <sup>[73]</sup>	80.5	97.1	99.6	78.6	63.6	62.8	80.4
DANN <sup>[11]</sup>	82.0	96.9	99.1	79.7	68.2	67.4	82.2
ADDA <sup>[110]</sup>	86.2	96.2	98.4	77.8	69.5	68.9	82.9
JAN <sup>[78]</sup>	85.4	97.4	99.8	84.7	68.6	70.0	84.3
GTA <sup>[18]</sup>	89.5	97.9	99.8	87.7	72.8	71.4	86.5
CAN <sup>[19]</sup>	81.5	98.2	99.7	85.5	65.9	63.4	82.4
JDDA <sup>[20]</sup>	82.6	95.2	99.7	79.8	57.4	66.7	80.2
MRAN <sup>[169]</sup>	91.4	96.9	99.8	86.4	68.3	70.9	85.6
MEDA <sup>[170]</sup>	86.2	97.2	99.4	85.3	72.4	74.0	85.8
MCD <sup>[171]</sup>	92.2	88.6	98.5	100.0	69.5	69.7	86.5
DSAN <sup>[159]</sup>	93.6	98.4	100.0	90.2	73.5	74.8	88.4
ALDA <sup>[172]</sup>	95.6	97.7	100.0	94.0	72.2	72.5	88.7
MDD <sup>[173]</sup>	93.5	94.5	98.4	100.0	74.6	72.2	88.9
DADA <sup>[174]</sup>	92.3	99.2	100.0	93.9	74.4	74.2	89.0
GVB-GD <sup>[175]</sup>	95.0	94.8	98.7	100.0	73.4	73.7	89.3
SPL <sup>[176]</sup>	92.7	98.7	99.8	93.0	76.4	76.8	89.6
GSDA <sup>[177]</sup>	95.7	99.1	100.0	94.8	73.5	74.9	89.7
SRDC <sup>[178]</sup>	95.7	99.2	100.0	95.8	76.7	77.1	90.8
CDAN <sup>[166]</sup>	93.1	98.2	100.0	89.8	70.1	68.0	86.6

表7 Office-Home数据集上的评价结果

Table 7 Evaluation results on Office-Home dataset

评价算法	Ar-Cl	Ar-Pr	Ar-Rw	Cl-Ar	Cl-Pr	Cl-Rw	Pr-Ar	Pr-Cl	Pr-Rw	Rw-Ar	Rw-Cl	Rw-Pr	平均
AlexNet <sup>[178]</sup>	26.4	32.6	41.3	22.1	41.7	42.1	20.5	20.3	51.1	31.0	27.9	54.9	34.3
ResNet-50 <sup>[168]</sup>	34.9	50.0	58.0	37.4	41.9	46.2	38.5	31.2	60.4	53.9	41.2	59.9	46.1
DAN <sup>[173]</sup>	43.6	57.0	67.9	45.8	56.5	60.4	44.0	43.6	67.7	63.1	51.5	74.3	56.3
DANN <sup>[11]</sup>	45.6	59.3	70.1	47.0	58.5	60.9	46.1	43.7	68.5	63.2	51.8	76.8	57.6
JAN <sup>[78]</sup>	45.9	61.2	68.9	50.4	59.7	61.0	45.8	43.4	70.3	63.9	52.4	76.8	58.3
MCD <sup>[171]</sup>	48.9	68.3	74.6	61.3	67.6	68.8	57	47.1	75.1	69.1	52.2	79.6	64.1
MEDA <sup>[170]</sup>	55.2	76.2	77.3	58.0	73.7	71.9	59.3	52.4	77.9	68.2	57.5	81.8	67.5
MRAN <sup>[169]</sup>	53.8	68.6	75.0	57.3	68.5	68.3	58.5	54.6	77.5	70.4	60.0	82.2	66.2
DSAN <sup>[159]</sup>	54.4	70.8	75.4	60.4	67.8	68.0	62.6	55.9	78.5	73.8	60.6	83.1	67.6
MDD <sup>[173]</sup>	54.9	73.7	77.8	60.0	71.4	71.8	61.2	53.6	78.1	72.5	60.2	82.3	68.1
DRMEA <sup>[179]</sup>	52.3	73.0	77.3	64.3	72.0	71.8	63.6	52.7	78.5	72.0	57.7	81.6	68.1
GSDA <sup>[176]</sup>	61.3	76.1	79.4	65.4	73.3	74.3	65.0	53.2	80.0	72.2	60.6	83.1	70.3
GVB-GD <sup>[174]</sup>	57.0	74.7	79.8	64.6	74.1	74.6	65.2	55.1	81.0	74.6	59.7	84.3	70.4
SRDC <sup>[177]</sup>	52.3	76.3	81.0	69.5	76.2	78.0	68.7	53.8	81.7	76.3	57.1	85.0	71.3
CDAN <sup>[166]</sup>	49.0	69.3	74.5	54.4	66.0	68.4	55.6	48.3	75.9	68.4	55.4	80.5	63.8

表8 深度域适应相关经典算法

Table 8 Implementation of classical algorithms about deep domain adaptation

算法名称	实现平台	文献发表年份
DaNN <sup>[71]</sup>	PyTorch	2014年
DDC <sup>[72]</sup>	PyTorch	2014年
DAN <sup>[73]</sup>	PyTorch	2015年
RTN <sup>[30]</sup>	Caffe	2016年
ADDA <sup>[110]</sup>	Tensorflow/Pytorch	2017年
SAN <sup>[111]</sup>	Caffe/Pytorch	2018年
CDCL <sup>[180]</sup>	Tensorflow	2020年
DSAN <sup>[164]</sup>	Pytorch	2021年
FixBi <sup>[117]</sup>	Pytorch	2021年

## 6 当前问题及未来研究方向

数据集的快速增长,为深度域适应发展提供了便捷,新方法层出不穷,实验结果更加精确。在此,本文展望未来的可研究方向,以进一步加强现有的工作。

### (1) 双向域适应

域适应的发展仍然面临更高难度挑战,例如,在域适应问题中,SVHN→MNIST的准确率能够从70.7%提高到99.3%<sup>[181]</sup>。而对于MNIST→SVHN的反向域适应,如果没有针对特定问题的超参数调优,准确率最高只能达到81.7%<sup>[182]</sup>。这表明反向域适应存在较大的难度,因此在未来的研究中,仍然需要加强对双向域适应问题的研究。

### (2) 超参数调优

反向验证和针对特定问题的匹配方法已经应用于超参数调优,这些方法并不需要目标标签。由于计算成本的增加<sup>[183]</sup>或反向验证精度与测试精度<sup>[184]</sup>不一致的问题存在,大多数方法并没有将反向验证广泛使用。Perone等<sup>[183]</sup>针对特定问题提出了匹配领域之间像素强度的方法,然而这种方法不具有普遍性。因此,超参数调优方法应该被开发应用于更广泛的问题,目前,这仍然是一个值得研究的开放领域。

### (3) 结合改进的图像到图像的迁移方法

结合改进的图像到图像的迁移方法也被应用在域适应中。例如,由于区域差异较大,Bousmalis等<sup>[122]</sup>的PixelDA方法难以直接应用。因此,类似于XGAN<sup>[123]</sup>(图像到图像的转换,可以支持差异较大的域转换)方法被开发出来。此类方法允许域之间存在更实质性的差异,并且可以直接扩展到域适应问题中。同样,类似StarGAN<sup>[185]</sup>的图像到图像的转换方法也已经被扩展应用到多域适应问题中。

### (4) 数据集的局限性

在不同情况和特定任务下,可以获得不同的源域和目标域数据,然而,域适应中数据集相比于深度学习中通用数据集较小。Sankaranarayanan等<sup>[186]</sup>注意到基于GAN网络模型的域适应方法需要大量的训练数据,如果源域或目标域数据集太小,实验结果会变差。而且,目前大部分领域适应数据集为计算机视觉方向。Nidadavolu等<sup>[187]</sup>认为,当前研究需要针对这种低资源情况进行调整。为了促进其他应用领域的研究,应该多创建其他领域相关数据集。

### (5) 异构数据挖掘

虽然深度域适应技术已被成功应用于许多实际应用中,如图像分类、风格转换、目标检测、人脸识别

别、语义分割和行人重识别等。但现有的算法大多集中于同构深度数据挖掘,即源域和目标域之间的特征空间是相同的。而这一假设在许多应用中并不成立,如图像域到文本域的迁移。我们希望在没有这种苛刻假设的情况下传递领域知识,更好地解决目标域任务。虽然目前已有异构域适应的相关研究被提出,但其实际效果并不佳,目前尚未有有效算法突破这一瓶颈。因此,异构深度数据挖掘在未来值得被进一步关注和研究。

#### (6) 探索域边界

现有的深度域适应研究方法通常会明确给定分布已知的源域和目标域。但在实际情景中,如何准确定义分布未知的域边界对域适应建模至关重要。例如,在自动驾驶场景中,当前驾驶状态受多个因素影响,如天气、光照、路面状况和车速等,而驾驶状态会随着客观因素的影响连续发生变化。此时面临的首要问题是对域边界进行检测,以判断何种驾驶状态下需要对模型进行迁移。遗憾的是,目前域边界探索研究还相对薄弱,将来有待更深入的研究。

#### (7) 最优纳什平衡点

基于对抗的域适应模型通过博弈的训练策略实现训练,领域生成器与域判别器之间需要达到最优纳什平衡点,以确保模型能够筛选和生成高质量样本<sup>[188]</sup>。但在实际场景中,生成对抗网络通常存在不同的平衡态,若模型训练未收敛到最优平衡点,则会导致模型学习不充分,即纳什平衡任务存在多个平衡点且多个平衡点之间关联性弱的问题。因此,如何寻找全部平衡点仍是目前学术界的公认难点之一,受该限制的影响,生成对抗机制在理论层面存在非最优平衡状态等诸多不利因素。遗憾的是,现有方法无法从理论证明角度出发推导模型的全局最优纳什平衡点,从而导致模型陷入局部平衡,限制模型实际表现。因此,这一方向仍然是未来面临的重要挑战。

## 7 结束语

针对基于深度学习的域适应问题,本文首先回顾了深度域适应近年来的研究现状,并依据是否涉及对抗训练机制,对现有相关工作进行了分析并将其分为非对抗域适应和对抗域适应两类。其次,当前现有的域适应方法在实际应用中仍然面临较大瓶颈,本文重点对深度域适应方法所面临的问题进行了总结,并指出未来可研究的方向。而部分域适应、开集域适应和通用域适应等相关方法,同样值得被深入探讨和研究。最后,希望本文工作能帮助读者了解该领域的研究现状,为读者相关研究提供借鉴。

### 参考文献:

- [1] BRUZZONE L, MARCONCINI M. Domain adaptation problems: A DASVM classification technique and a circular validation strategy[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 32(5): 770-787.
- [2] CHU W S, DE LA TORRE F, COHN J F. Selective transfer machine for personalized facial action unit detection[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Portland: IEEE, 2013: 3515-3522.
- [3] GONG B, GRAUMAN K, SHA F. Connecting the dots with landmarks: Discriminatively learning domain-invariant features for unsupervised domain adaptation[C]// *Proceedings of International Conference on Machine Learning*. Atlanta: ACM, 2013: 222-230.
- [4] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis[J]. *IEEE Transactions on Neural Networks*, 2010, 22(2): 199-210.
- [5] 张文婧,宋鹏,陈栋梁,等.基于稀疏子空间迁移学习的跨域人脸表情识别[J]. *数据采集与处理*, 2021, 36(1): 113-121.  
ZHANG Wenjing, SONG Peng, CHEN Dongliang, et al. Cross-domain facial expression recognition based on sparse subspace transfer learning[J]. *Journal of Data Acquisition and Processing*, 2021, 36(1): 113-121.
- [6] PACHORI S, DESHPANDE A, RAMAN S. Hashing in the zero shot framework with domain adaptation[J]. *Neurocomputing*, 2018, 275(2): 2137-2149.

- [7] CAO Y, LONG M, WANG J. Unsupervised domain adaptation with distribution matching machines[C]//Proceedings of the 32nd Thirty-Second AAAI Conference on Artificial Intelligence. Louisiana: AAAI, 2018: 16571-16580.
- [8] 付玉香,秦永斌,申国伟.基于迁移学习的敏感数据隐私保护方法[J].数据采集与处理,2019,34(3):422-431.  
FU Yuxiang, QIN Yongbin, SHEN Guowei. Sensitive data privacy protection method based on transfer learning[J]. Journal of Data Acquisition and Processing, 2019, 34(3): 422-431.
- [9] ZHAO Z, CHEN Y, LIU J, et al. Cross - people mobile - phone based activity recognition[C]//Proceedings of the 22nd International Joint Conference on Artificial Intelligence. Barcelona Catalonia: AAAI, 2011: 3731-3738.
- [10] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J/OL]. Advances in Neural Information Processing Systems. (2014-06-10) [2021-05-30]. <https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>.
- [11] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//Proceedings of International Conference on Machine Learning. Lille: ACM, 2015: 1180-1189.
- [12] SHAO L, ZHU F, LI X. Transfer learning for visual categorization: A survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 26(5): 1019-1034.
- [13] DAY O, KHOSHGOFTAAR T M. A survey on heterogeneous transfer learning[J]. Journal of Big Data, 2017, 4(1): 29-49.
- [14] CSURKA G. Domain adaptation for visual applications: A comprehensive survey[EB/OL]. (2017-02-17)[2021-05-30]. <https://arxiv.org/pdf/1702.05374.pdf>.
- [15] WANG M, DENG W. Deep visual domain adaptation: A survey[J]. Neurocomputing, 2018, 312: 135-153.
- [16] WILSON G, COOK D J. A survey of unsupervised deep domain adaptation[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2020, 11(5): 1-46.
- [17] ZHUANG F, QI Z, DUAN K, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2020, 109(1): 43-76.
- [18] SANKARANARAYANAN S, BALAJI Y, CASTILLO C D, et al. Generate to adapt: Aligning domains using generative adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Utah: IEEE, 2018: 8503-8512.
- [19] ZHANG W, OUYANG W, LI W, et al. Collaborative and adversarial network for unsupervised domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Utah: IEEE, 2018: 3801-3809.
- [20] CHEN C, CHEN Z, JIANG B, et al. Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Hawaii: AAAI, 2019: 3296-3303.
- [21] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 22(10): 1345-1359.
- [22] LIU F, ZHANG G, LU J. Heterogeneous domain adaptation: An unsupervised approach[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(12): 5588-5602.
- [23] PANAREDA BUSTO P, GALL J. Open set domain adaptation[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 754-763.
- [24] YOU K, LONG M, CAO Z, et al. Universal domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 2720-2729.
- [25] LIU X, GUO Z, LI S, et al. Adversarial unsupervised domain adaptation with conditional and label shift: Infer, Align and Iterate[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 10367-10376.
- [26] SAHA A, RAI P, DAUMÉ H, et al. Active supervised domain adaptation[C]//Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Berlin: Springer, 2011: 97-112.
- [27] HE G, LIU X, FAN F, et al. Classification-aware semi-supervised domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020: 964-965.
- [28] AO S, LI X, LING C. Fast generalized distillation for semi-supervised domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. San Francisco: AAAI, 2017: 31-33.
- [29] SAITO K, KIM D, SCLAROFF S, et al. Semi-supervised domain adaptation via minimax entropy[C]//Proceedings of the

- IEEE International Conference on Computer Vision. Long Beach: IEEE, 2019: 8050-8058.
- [30] LONG M, ZHU H, WANG J, et al. Unsupervised domain adaptation with residual transfer networks[C]//Proceedings of Advances in Neural Information Processing Systems. Barcelona: MIT Press, 2016: 136-144.
- [31] LONG M, WANG J, DING G, et al. Transfer joint matching for unsupervised domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 1410-1417.
- [32] VOLPI R, MORERIO P, SAVARESE S, et al. Adversarial feature augmentation for unsupervised domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 5495-5504.
- [33] WANG C, MAHADEVAN S. Heterogeneous domain adaptation using manifold alignment[C]//Proceedings of International Joint Conference on Artificial Intelligence. Barcelona: Morgan Kaufmann, 2011: 1541.
- [34] ZHOU J T, TSANG I W, PAN S J, et al. Heterogeneous domain adaptation for multiple classes[C]//Proceedings of Artificial Intelligence and Statistics. Reykjavic: Morgan kaufmann, 2014: 1095-1103.
- [35] LI J, LU K, HUANG Z, et al. Heterogeneous domain adaptation through progressive alignment[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 30(5): 1381-1391.
- [36] YANG T, HU L, SHI C, et al. HGAT: Heterogeneous graph attention networks for semi-supervised short text classification [J]. ACM Transactions on Information Systems (TOIS), 2021, 39(3): 1-29.
- [37] SHI W, GUPTE M S, WANG M D. Learning from heterogeneous data via contrastive learning: an application in multi-source COVID-19 radiography[C]//Proceedings of 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). [S.l.]: IEEE, 2021: 1-4.
- [38] TAN B, SONG Y, ZHONG E, et al. Transitive transfer learning[C]//Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2015: 1155-1164.
- [39] TAN B, ZHANG Y, PAN S J, et al. Distant domain transfer learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. San Francisco: AAAI, 2017: 301-302.
- [40] MOON J H, DAS D, LEE C S G. Multi-step online unsupervised domain adaptation[C]//Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona: IEEE, 2020: 41172-41576.
- [41] CUI S, WANG S, ZHUO J, et al. Gradually vanishing bridge for adversarial domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 12455-12464.
- [42] MANCINI M, BULO S R, CAPUTO B, et al. Adagraph: Unifying predictive and continuous domain adaptation through graphs[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2019: 6568-6577.
- [43] MA A, YOU F, JING M, et al. Multi-source domain adaptation with graph embedding and adaptive label prediction[J]. Information Processing & Management, 2020, 57(6): 102367.
- [44] MORSING L H, SHEIKH - OMAR O A, IOSIFIDIS A. Supervised domain adaptation using graph embedding[C]//Proceedings of 2020 25th International Conference on Pattern Recognition (ICPR). Milan: IEEE, 2021: 7841-7847.
- [45] TZENG E, HOFFMAN J, DARRELL T, et al. Simultaneous deep transfer across domains and tasks[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 4068-4076.
- [46] TIMNIT G, JUDY H, LI F. Fine-grained recognition in the wild: A multi-task domain adaptation approach[C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul: IEEE, 2017: 1349-1358.
- [47] YAN H, DING Y, LI P, et al. Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2272-2281.
- [48] YU Q, HASHIMOTO A, USHIKU Y. Divergence optimization for noisy universal domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: USA, IEEE, 2021: 2515-2524.
- [49] ROZANTSEV A, SALZMANN M, FUA P. Beyond sharing weights for deep domain adaptation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(4): 801-814.
- [50] LI Y, WANG N, SHI J, et al. Revisiting batch normalization for practical domain adaptation[EB/OL]. (2016-03-15)[2021-05-30]. <https://arxiv.org/pdf/1603.04779.pdf?ref=https://githubhelp.com>.

- [51] KULLBACK S, LEIBLER R A. On information and sufficiency[J]. *The Annals of Mathematical Statistics*, 1951, 22(1): 79-86.
- [52] FUGLEDE B, TOPSOE F. Jensen - Shannon divergence and Hilbert space embedding[C]//*Proceedings of International Symposium on Information Theory*. Chicago: IEEE, 2004: 31-33.
- [53] WANG J, CHEN Y, HU L, et al. Stratified transfer learning for cross-domain activity recognition[C]//*Proceedings of 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. Athens: IEEE, 2018: 1-10.
- [54] GRETTON A, BOUSQUET O, SMOLA A, et al. Measuring statistical dependence with Hilbert - Schmidt norms[C]//*International Conference on Algorithmic Learning Theory*. Singapore: Springer, 2005: 63-77.
- [55] BORGWARDT K M, GRETTON A, RASCH M J, et al. Integrating structured biological data by kernel maximum mean discrepancy[J]. *Bioinformatics*, 2006, 22(14): 49-57.
- [56] SUN H. Mercer theorem for RKHS on noncompact sets[J]. *Journal of Complexity*, 2005, 21(3): 337-349.
- [57] FU J, WU X, ZHANG S, et al. Improved open set domain adaptation with backpropagation[C]//*Proceedings of 2019 IEEE International Conference on Image Processing (ICIP)*. [S.l.]: IEEE, 2019: 2506-2510.
- [58] DAI W, YANG Q, XUE G R, et al. Self-taught clustering[C]//*Proceedings of the 25th International Conference on Machine learning*. Helsinki: ACM, 2008: 200-207.
- [59] DAVIS J, DOMINGOS P. Deep transfer via second - order Markov logi[C]//*Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal: ACM, 2009: 217-224.
- [60] GILES J, ANG K K, MIHAYLOVA L S, et al. A Subject-to-subject transfer learning framework based on jensen-shannon divergence for improving brain - computer interface[C]//*Proceedings of 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton: IEEE, 2019: 3087-3091.
- [61] CHEN W H, CHO P C, JIANG Y L. Activity recognition using transfer learning[J]. *Sens Mater*, 2017, 29(7): 897-904.
- [62] DEY S, MADIKERI S, MOTLICEK P, et al. Information theoretic clustering for unsupervised domain - adaptation[C]//*Proceedings of 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai: IEEE, 2016: 5580-5584.
- [63] ZHANG Y, WANG N, CAI S. Adversarial sliced Wasserstein domain adaptation networks[J]. *Image and Vision Computing*, 2020, 102: 103974.
- [64] LEE C Y, BATRA T, BAIG M H, et al. Sliced wasserstein discrepancy for unsupervised domain adaptation[C]//*Proceedings of Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2019: 10285-10295.
- [65] CAI G, LIN Q, CHEN N. Cross-domain text sentiment classification based on Wasserstein distance[C]//*Proceedings of International Conference on Security with Intelligent Computing and Big-data Services*. Guilin: Springer, 2018: 280-291.
- [66] WANG H Y, YANG Q. Transfer learning by structural analogy[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. San Francisco: AAAI, 2011: 57- 58.
- [67] XIAO M, GUO Y. Feature space independent semi-supervised domain adaptation via kernel matching[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 37(1): 54-66.
- [68] YAN K, KOU L, ZHANG D. Learning domain-invariant subspace using domain features and independence maximization[J]. *IEEE Transactions on Cybernetics*, 2017, 48(1): 288-299.
- [69] LONG M, WANG J, DING G, et al. Adaptation regularization: A general framework for transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2013, 26(5): 1076-1089.
- [70] LONG M, WANG J, DING G, et al. Transfer feature learning with joint distribution adaptation[C]//*Proceedings of the IEEE International Conference on Computer Vision*. Sydney: IEEE, 2013: 2200-2207.
- [71] RÜSCHENDORF L. The Wasserstein distance and approximation theorems[J]. *Probability Theory and Related Fields*, 1985, 70(1): 117-129.
- [72] GHIFARY M, KLEIJN W B, ZHANG M. Domain adaptive neural networks for object recognition[C]//*Proceedings of Pacific Rim International Conference on Artificial Intelligence*. Queensland: Springer, 2014: 898-904.
- [73] TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: Maximizing for domain invariance[EB/OL]. (2014-07-07)[2021-05-30]. <https://arxiv.org/pdf/1412.3474.pdf>.



- [74] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks[C]//Proceedings of International Conference on Machine Learning. Lille: ACM, 2015: 97-105.
- [75] DU Z, LI J, SU H, et al. Cross-domain gradient discrepancy minimization for unsupervised domain adaptation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 3937-3946.
- [76] SAITO K, USHIKU Y, HARADA T. Asymmetric tri-training for unsupervised domain adaptation[C]//Proceedings of International Conference on Machine Learning. Sydney: [s.n.], 2017: 2988-2997.
- [77] ZHU Y, ZHUANG F, WANG J, et al. Deep subdomain adaptation network for image classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 32(4): 1713-1722.
- [78] LONG M, ZHU H, WANG J, et al. Deep transfer learning with joint adaptation networks[C]//Proceedings of International Conference on Machine Learning. Sydney: ACM, 2017: 2208-2217.
- [79] ZHANG X, YU F X, CHANG S F, et al. Deep transfer network: Unsupervised domain adaptation[EB/OL]. (2015-03-02) [2021-05-30]. <https://arxiv.org/pdf/1503.00591.pdf>.
- [80] 马闯, 田青, 孙赫阳, 等. 基于双重权重偏差建模的无监督域适应[J]. *计算机科学*, 2021, 48(2): 217-223.  
MA Chuang, TIAN Qing, SUN Heyang, et al. Unsupervised domain adaptation based on weighting dual biases[J]. *Computer Science*, 2021, 48(2): 217-223.
- [81] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis[J]. *IEEE Transactions on Neural Networks*, 2010, 22(2): 199-210.
- [82] DORRI F, GHODSI A. Adapting component analysis[C]//Proceedings of 2012 IEEE 12th International Conference on Data Mining. Brussels: IEEE, 2012: 846-851.
- [83] DUAN L, TSANG I W, XU D. Domain transfer multiple kernel learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(3): 465-479.
- [84] ZHANG K, SCHÖLKOPF B, MUANDET K, et al. Domain adaptation under target and conditional shift[C]//Proceedings of International Conference on Machine Learning. Atlanta: ACM, 2013: 819-827.
- [85] GONG M, ZHANG K, LIU T, et al. Domain adaptation with conditional transferable components[C]//Proceedings of International Conference on Machine Learning. New York: ACM, 2016: 2839-2848.
- [86] ZHANG J, LI W, OGUNBONA P, et al. Joint geometrical and statistical alignment for visual domain adaptation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 1859-1867.
- [87] LONG M, ZHU H, WANG J, et al. Deep transfer learning with joint adaptation networks[C]// International Conference on Machine Learning. Sydney: ACM, 2017: 2208-2217.
- [88] WANG J, CHEN Y, HAO S, et al. Balanced distribution adaptation for transfer learning[C]//Proceedings of 2017 IEEE International Conference on Data Mining (ICDM). Honolulu: IEEE, 2017: 1129-1134.
- [89] LI J, JING M, LU K, et al. Locality preserving joint transfer for domain adaptation[J]. *IEEE Transactions on Image Processing*, 2019, 28(12): 6103-6115.
- [90] HINTON G E, ZEMEL R S. Autoencoders, minimum description length, and Helmholtz free energy[J/OL]. *Advances in Neural Information Processing Systems*, 1993[2022-05-30]. <https://proceedings.neurips.cc/paper/1993>.
- [91] VINCENT P, LAROCHELLE H, LAJOIE I, et al. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion[J]. *Journal of Machine Learning Research*, 2010, 11(12): 3371-3408.
- [92] CHEN M, XU Z E, WEINBERGER K Q, et al. Marginalized denoising autoencoders for domain adaptation[C]// Proceedings of International Conference on Machine Learning. Edinburgh: [s.n.], 2012.
- [93] GHIFARY M, KLEIJN W B, ZHANG M, et al. Deep reconstruction-classification networks for unsupervised domain adaptation[C]//Proceedings of European Conference on Computer Vision. Amsterdam: Springer, 2016: 597-613.
- [94] BOUSMALIS K, TRIGEORGIS G, SILBERMAN N, et al. Domain separation networks[C]//Proceedings of Advances in Neural Information Processing Systems. Barcelona: MIT Press, 2016: 343-351.
- [95] TSAI J C, CHIEN J T. Adversarial domain separation and adaptation[C]//Proceedings of 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP). Tokyo: IEEE, 2017: 1-6.
- [96] ZHUANG F, CHENG X, LUO P, et al. Supervised representation learning: Transfer learning with deep autoencoders[C]//

- Proceedings of the 24th International Joint Conference on Artificial Intelligence. Austin: AAAI, 2015: 99-107.
- [97] YU Y, GONG Z, ZHONG P, et al. Unsupervised representation learning with deep convolutional neural network for remote sensing images[C]//Proceedings of International Conference on Image and Graphics. Shanghai: Springer, 2017: 97-108.
- [98] GAUTHIER J. Conditional generative adversarial nets for convolutional face generation[J]. Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, 2014, 2014(5): 2-17.
- [99] CHEN X, DUAN Y, HOUTHOOFT R, et al. Infogan: Interpretable Representation learning by information maximizing generative adversarial nets[C]//Proceedings of Advances in Neural Information Processing Systems. Barcelona: MIT Press, 2016: 2172-2180.
- [100] ODENA A, OLAH C, SHLENS J. Conditional image synthesis with auxiliary classifier gans[C]//Proceedings of International Conference on Machine Learning. Sydney: ACM, 2017: 2642-2651.
- [101] CRESWELL A, BHARATH A A. Denoising adversarial autoencoders[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 30(4): 968-984.
- [102] LI H, PAN S J, WANG S, et al. Domain generalization with adversarial feature learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 5400-5409.
- [103] DU C, DU C, XIE X, et al. Multi-view adversarially learned inference for cross-domain joint distribution matching[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1348-1357.
- [104] METZ L, POOLE B, PFAU D, et al. Unrolled generative adversarial networks[EB/OL].(2016-11-07)[2021-05-30].<https://arxiv.org/pdf/1611.02163.pdf>.
- [105] NOWOZIN S, CSEKE B, TOMIOKA R. F-gan: Training generative neural samplers using variational divergence minimization[C]// Advances in Neural Information Processing Systems. Barcelona: MIT Press, 2016: 271-279.
- [106] MAO X, LI Q, XIE H, et al. On the effectiveness of least squares generative adversarial networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(12): 2947-2960.
- [107] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//Proceedings of International Conference on Machine Learning. Sydney: ACM, 2017: 214-223.
- [108] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[C]//Proceedings of Advances in Neural Information Processing Systems. Long Beach: MIT Press, 2017: 5767-5777.
- [109] PETZKA H, FISCHER A, LUKOVNICOV D. On the regularization of wasserstein gans[EB/OL].(2018-04-09)[2021-05-30]. <https://arxiv.org/pdf/1709.08894.pdf>.
- [110] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 7167-7176.
- [111] CAO Z, LONG M, WANG J, et al. Partial transfer learning with selective adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 2724-2732.
- [112] ZHANG Y, DAVISON B D. Adversarial consistent learning on partial domain adaptation of plantclef 2020 challenge[EB/OL].(2020-09-19)[2021-05-30]. <https://arxiv.org/pdf/2009.09289.pdf>.
- [113] CHEN Z, CHEN C, CHENG Z, et al. Selective transfer with reinforced transfer network for partial domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 12706-12714.
- [114] XU M, ZHANG J, NI B, et al. Adversarial domain adaptation with domain mixup[C]// Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34(4): 6502-6509.
- [115] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2223-2232.
- [116] CHEN Y C, LIN Y Y, YANG M H, et al. Crdoco: Pixel-level domain transfer with cross-domain consistency[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 1791-1800.
- [117] NA J, JUNG H, CHANG H J, et al. FixBi: Bridging domain spaces for unsupervised domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 1094-1103.
- [118] CAO Y, LIU B, LONG M, et al. Hashgan: Deep learning to hash with pair conditional wasserstein gan[C]//Proceedings of

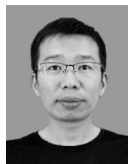
- the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 1287-1296.
- [119] SHEN J, QU Y, ZHANG W, et al. Wasserstein distance guided representation learning for domain adaptation[C]// Proceedings of AAAI Conference on Artificial Intelligence. New Orleans: [s.n.], 2018.
- [120] 蔡德润, 李红艳. 基于度量学习的无监督域适应方法及其在死亡风险预测上的应用[J]. 计算机研究与发展, 2022, 59(3): 674-682.
- CAI Derun, LI Hongyan. Unsupervised domain adaptation based on measurement learning and its application to mortality risk prediction[J]. Journal of Computer Research and Development, 2022, 59(3): 674-682.
- [121] RAHMAN M M, FOOKES C, BAKTASHMOTLAGH M, et al. Correlation-aware adversarial domain adaptation and generalization[J]. Pattern Recognition, 2020. DOI:10.1016/j.patcog.2019.107124.
- [122] BOUSMALIS K, SILBERMAN N, DOHAN D, et al. Unsupervised pixel-level domain adaptation with generative adversarial networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 3722-3731.
- [123] ROYER A, BOUSMALIS K, GOUWS S, et al. Xgan: Unsupervised image-to-image translation for many-to-many mappings [C]// Proceedings of Domain Adaptation for Visual Understanding. Seattle: IEEE, 2020: 33-49.
- [124] CUI S, WANG S, ZHUO J, et al. Gradually vanishing bridge for adversarial domain adaptation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 12455-12464.
- [125] AHMED S M, RAYCHAUDHUR D S, PAUL S, et al. Unsupervised Multi-source domain adaptation without access to source data[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 10103-10112.
- [126] LIN J, YUAN K, WARD R, et al. Xnet: Task-specific attentional domain adaptation for satellite-to-aerial scene[J]. Neurocomputing, 2020, 406: 215-223.
- [127] SALI A W, JIANG J, EGNER T. Neural mechanisms of strategic adaptation in attentional flexibility[J]. Journal of Cognitive Neuroscience, 2020, 32(5): 989-1008.
- [128] CHEN C, DOU Q, CHEN H, et al. Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest X-ray segmentation[C]// Proceedings of International Workshop on Machine Learning in Medical Imaging. Stockholm: ACM, 2018: 143-151.
- [129] ZHANG Y, MIAO S, MANSI T, et al. Task driven generative modeling for unsupervised domain adaptation: Application to X-ray image segmentation[C]// Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada: Springer, 2018: 599-607.
- [130] CHEN Y, LI W, SAKARIDIS C, et al. Domain adaptive faster F-CNN for object detection in the wild[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 3339-3348.
- [131] LI D, HUANG J B, LI Y, et al. Weakly supervised object localization with progressive domain adaptation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 3512-3520.
- [132] INOUE N, FURUTA R, YAMASAKI T, et al. Cross-domain weakly-supervised object detection through progressive domain adaptation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 5001-5009.
- [133] BAK S, CARR P, LALONDE J F, et al. Domain adaptation through synthesis for unsupervised person re-identification[C]// Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 189-205.
- [134] DENG W, ZHENG L, YE Q, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 994-1003.
- [135] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17(1): 2096-2130.
- [136] LI Y J, YANG F E, LIU Y C, et al. Adaptation and re-identification network: An unsupervised deep transfer learning approach to person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake: IEEE, 2018: 172-178.

- [137] WEI L, ZHANG S, GAO W, et al. Person transfer gan to bridge domain gap for person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 79-88.
- [138] ZHONG Z, ZHENG L, LUO Z, et al. Invariance matters: Exemplar memory for domain adaptive person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 598-607.
- [139] ZHONG Z, ZHENG L, ZHENG Z, et al. Camera style adaptation for person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2018: 5157-5166.
- [140] ZHANG Y, BARZILAY R, JAAKKOLA T, et al. Aspect-augmented adversarial networks for domain adaptation[J]. Transactions of the Association for Computational Linguistics, 2017, 5(2): 515-528.
- [141] LIU K, ZHAO J, et al. Domain adaptation in NLP based on hybrid generative and discriminative model[C]//Proceedings of 2008 Chinese Conference on Pattern Recognition. Beijing: IEEE, 2008: 1-6.
- [142] GUO H, PASUNURU R, BANSAL M, et al. Multi-source domain adaptation for text classification via distancenet-bandits [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 7830-7838.
- [143] YANG Z, HU J, SALAKHUTDINOV R, et al. Semi-supervised QA with generative domain-adaptive nets[EB/OL].(2017-04-22)[2021-05-30]. <https://arxiv.org/pdf/1702.02206.pdf>.
- [144] BRITZ D, LE Q, PRYZANT R, et al. Effective domain mixing for neural machine translation[C]// Proceedings of the Second Conference on Machine Translation. Copenhagen: IEEE, 2017: 118-126.
- [145] WANG R, UTIYAMA M, FINCH A, et al. Sentence selection and weighting for neural machine translation domain adaptation[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018, 26(10): 1727-1741.
- [146] CHU C, DABRE R, KUROHASHI S, et al. A comprehensive empirical comparison of domain adaptation methods for neural machine translation[J]. Journal of Information Processing, 2018, 26(9): 529-538.
- [147] TOBORE I, LI J, YUHANG L, et al. Deep learning intervention for health care challenges: Some biomedical domain considerations[J]. JMIR mHealth and uHealth, 2019, 7(8): 11966.
- [148] SUN S, ZHANG B, XIE L, et al. An unsupervised deep domain adaptation approach for robust speech recognition[J]. Neurocomputing, 2017, 257: 79-87.
- [149] TSAI Y H, HUNG W C, SCHULTER S, et al. Learning to adapt structured output space for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 7472-7481.
- [150] KUMAGAI A, IWATA T, FUJIWARA Y. Transfer anomaly detection by inferring latent domain representations[J/OL]. Advances in Neural Information Processing Systems, 2019[2021-05-30]. <https://proceedings.nips.cc/paper/2019/hash/7895fc13088ee37f511913bac71fa66f-Abstract.html>.
- [151] ZHAO M, YUE S, KATABI D, et al. Learning sleep stages from radio signals: A conditional adversarial architecture[C]// Proceedings of International Conference on Machine Learning. Sydney: ACM, 2017: 4100-4109.
- [152] FAWAZ H I, FORESTIER G, WEBER J, et al. Transfer learning for time series classification[C]//Proceedings of 2018 IEEE International Conference on Big Data (Big Data). Seattle: IEEE, 2018: 1367-1376.
- [153] VESAL S, RAVIKUMAR N, MAIER A. Automated multi-sequence cardiac MRI segmentation using supervised domain adaptation[C]//Proceedings of International Workshop on Statistical Atlases and Computational Models of the Heart. Shenzhen: Springer, 2019: 300-308.
- [154] ORBES-ARTEAGA M, VARSAVSKY T, SUDRE C H, et al. Multi-domain adaptation in brain MRI through paired consistency and adversarial learning[C]//Proceedings of Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data. Shenzhen: Springer, 2019: 54-62.
- [155] LI C, WANG G, DE MELO G. Context-based few-shot word representation learning[C]//Proceedings of 2018 IEEE 12th International Conference on Semantic Computing (ICSC). Laguna Hills: IEEE, 2018: 239-242.
- [156] LIU C Y, SUN W B, CHAO W H, et al. Convolution neural network for relation extraction[C]//Proceedings of International Conference on Advanced Data Mining and Applications. Sydney: Springer, 2013: 231-242.
- [157] BLANCHARD G, LEE G, SCOTT C, et al. Generalizing from several related classification tasks to a new unlabeled sample [C]//Proceedings of Advances in Neural Information Processing Systems. Granada: MIT Press, 2011: 2178-2186.
- [158] MUANDET K, BALDUZZI D, SCHÖLKOPF B, et al. Domain generalization via invariant feature presentation[C]//

- Proceedings of International Conference on Machine Learning. Atlanta: ACM, 2013: 10-18.
- [159] GHIFARY M, BASTIAAN KLEIJN W, ZHANG M, et al. Domain generalization for object recognition with multi-task autoencoders[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 2551-2559.
- [160] CARLUCCI F M, RUSSO P, TOMMASI T, et al. Agnostic domain generalization[EB/OL]. (2018-08-30)[2021-05-30].<http://arXiv.preprintarXiv.1808.01102>, 2018.
- [161] AKUZAWA K, IWASAWA Y, MATSUO Y, et al. Adversarial invariant feature learning with accuracy constraint for domain generalization[C]//Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases.[S.l.]: Springer, 2019: 315-331.
- [162] COHEN G, AFSHAR S, TAPSON J, et al. EMNIST: Extending MNIST to handwritten letters[C]//Proceedings of 2017 International Joint Conference on Neural Networks (IJCNN). Alaska: IEEE, 2017: 2921-2926.
- [163] MOISEEV B, KONEV A, CHIGORIN A, et al. Evaluation of traffic sign recognition methods trained on synthetically generated data[C]//Proceedings of International Conference on Advanced Concepts for Intelligent Vision Systems. Auckland: Springer, 2013: 576-583.
- [164] STALLKAMP J, SCHLIPSING M, SALMEN J, et al. The German traffic sign recognition benchmark: A multi-class classification competition[C]//Proceedings of The 2011 International Joint Conference on Neural Networks. California: IEEE, 2011: 1453-1460.
- [165] SAENKO K, KULIS B, FRITZ M, et al. Adapting visual category models to new domains[C]//Proceedings of European Conference on Computer Vision. Heraklion: Springer, 2010: 213-226.
- [166] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation[J]. *Advances in neural information processing systems*, 2018.
- [167] PENG X, BAI Q, XIA X, et al. Moment matching for multi-source domain adaptation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 1406-1415.
- [168] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [169] ZHU Y, ZHUANG F, WANG J, et al. Multi-representation adaptation network for cross-domain image classification[J]. *Neural Networks*, 2019, 119: 214-221.
- [170] WANG J, FENG W, CHEN Y, et al. Visual domain adaptation with manifold embedded distribution alignment[C]//Proceedings of the 26th ACM International Conference on Multimedia. New York: ACM, 2018: 402-410.
- [171] SAITO K, WATANABE K, USHIKU Y, et al. Maximum classifier discrepancy for unsupervised domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 3723-3732.
- [172] CHEN M, ZHAO S, LIU H, et al. Adversarial-learned loss for domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34(4): 3521-3528.
- [173] ZHANG Y, LIU T, LONG M, et al. Bridging theory and algorithm for domain adaptation[C]//Proceedings of International Conference on Machine Learning. Long Beach: [s.n.], 2019: 7404-7413.
- [174] TANG H, JIA K. Discriminative adversarial domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34(4): 5940-5947.
- [175] WANG Q, BRECKON T. Unsupervised domain adaptation via structured prediction based selective pseudo-labeling[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York : AAAI, 2020, 34(4): 6243-6250.
- [176] HU L, KAN M, SHAN S, et al. Unsupervised domain adaptation with hierarchical gradient synchronization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 4043-4052.
- [177] TANG H, CHEN K, JIA K. Unsupervised domain adaptation via structurally regularized deep clustering[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 8725-8735.
- [178] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J/OL]. *Advances in Neural Information Processing Systems*, 2012[2021-05-30]. <https://proceedings.neurips.cc/paper/2012>.
- [179] LUO Y W, REN C X, GE P, et al. Unsupervised domain adaptation via discriminative manifold embedding and alignment [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34(4): 5029-5036.

- [180] LIN K, WANG L, LUO K, et al. Cross-domain complementary learning using pose for multi-person part segmentation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(3): 1066-1078.
- [181] HOFFMAN J, TZENG E, PARK T, et al. Cycada: Cycle-consistent adversarial domain adaptation[C]//Proceedings of International Conference on Machine Learning. Stockholm: ACM, 2018: 1989-1998.
- [182] KUMAR A, SATTIGERI P, WADHAWAN K, et al. Co-regularized alignment for unsupervised domain adaptation[J]. Advances in Neural Information Processing Systems, 2018, 31: 9367-9378.
- [183] PERONE C S, BALLESTER P, BARROS R C, et al. Unsupervised domain adaptation for medical imaging segmentation with self-ensembling[J]. NeuroImage, 2019, 194(26): 1-11.
- [184] BOUSMALIS K, TRIGEORGIS G, SILBERMAN N, et al. Domain separation networks[J/OL]. Advances in Neural Information Processing Systems, 2016[2021-05-30]. <https://proceedings.neurips.cc/paper/2016>.
- [185] CHOI Y, CHOI M, KIM M, et al. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 8789-8797.
- [186] SANKARANARAYANAN S, BALAJI Y, CASTILLO C D, et al. Generate to adapt: Aligning domains using generative adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 8503-8512.
- [187] NIDADAVOLU P S, KATARIA S, VILLALBA J, et al. Low-resource domain adaptation for speaker recognition using cycle-gans [C]//Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU). Sentosa: IEEE, 2019: 710-717.
- [188] FEDUS W, ROSCA M, LAKSHMINARAYANAN B, et al. Many paths to equilibrium: GANs do not need to decrease a divergence at every step[C]//Proceedings of International Conference on Learning Representations. Vancouver: [s.n.], 2018.

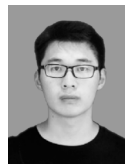
## 作者简介:



田青(1983-),通信作者,男,博士,副教授,硕士研究生导师,研究方向:机器学习、模式识别, E-mail: tianqing@nuist.edu.cn。



朱雅楠(1997-),女,硕士研究生,研究方向:机器学习、模式识别, E-mail: 20201220059@nuist.edu.cn。



马闾(1996-),男,硕士研究生,研究方向:机器学习、模式识别, E-mail: mac-boo@nuist.edu.cn。

(编辑:夏道家)